Assessing Time Series Forecast Models for Mixed Migration Flows in Europe

Mebelli Vasiliki

SID:

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

JANUARY 2021

THESSALONIKI – GREECE

Assessing Time Series Forecast Models for Mixed Migration Flows in Europe

Mebelli Vasiliki

SID:

Supervisor: Prof. Maria Drakaki

SCHOOL OF SCIENCE & TECHNOLOGY

A thesis submitted for the degree of

Master of Science (MSc) in Data Science

JANUARY 2021

THESSALONIKI – GREECE

Ackowledgment

Throughout the completion of the hereby dissertation thesis, Professor Maria Drakaki, has been a supportive supervisor. Her suggestions and feedback were so much appreciated and I would like to thank her for her overall contribution.

Abstract

This dissertation was written as a part of the MSc in Data Science at the International Hellenic University. It outlines the findings of applying 4 time series forecasting models (Exponential Smoothing Models, ARIMA, Neural Networks and Bayesian Models) on 3 different datasets. 3 European countries were selected to reflect on the ability of those models to predict the future behavior of an extremely volatile migrant mobility. Monthly records of first-time asylum-seekers were used from January of 2008 up to September of 2020 from the aforementioned countries and the results demonstrate clearly that there is still room for improvement on this research field.

Vasiliki Mebelli

4 January 2020

Contents

1 Introduction

Global migration is a highly complex dynamic system. Considering its importance, the scientific research towards the decipherment of its drivers has been scantily explored. The contemporary type of migration that has been in the spotlight since 2014 is mixed migration. The characteristics that make this type of migration so distinctive is the fact that from the beginning it has been a volatile international phenomenon with debatable impacts on society and economy. The mixed migration category consists of various types of migrant mobility. This thesis is focused on the forecasting methods of the least stable and predictable type of mixed migration category; asylum-seekers. In 2015, the first endeavor for proactive data-driven decision-making concerning refugees was the Winter Cell (UNHCR, 2015), which focused on estimating the movements of refugees, particularly approaching winter months, reaching Europe through Turkey. It was intended to last 4 months but eventually was extended due to the increased cases of such mobility. The interdisciplinarity of the team made possible for a model to be created that uses meteorological data to forecast migrant mobility to Greece. Since then, organizations that support similar purposes have been created, various projects have evolved and a more data-driven approach has been applied. One of the biggest challenges in forecast is the accurate measurement of uncertainty. Migration constitutes the most unpredictable demographic change. On top of that, asylum-related migration constitutes the most unstable and sensitive type of migration. Hence, asylum-related migration is almost impossible to predict. Fear of persecution can be a determinant factor to drive asylum-seekers to leave a country. Nevertheless, access to information contributes greatly to the selection of a final destination country. A promising theory which understands and, potentially, accurately forecasts asylum-related migration flows needs to include both “push-pull” factors. This thesis is an endeavor of assessing the accuracy and the performance of the state-of-the-art time series forecasting models in regards to asylum-related migration to Europe. At this point, it should be clarified that machine learning methods and statistical methods are different approaches with similar goals and will be combined throughout the thesis.

The paper is structured in IX sections. In Section I, some introductory key notes are mentioned. Then, in Section II the approved terminology along with the multidimensionality of mixed migration will be presented. Next, in Section III there will be a review of the literature relevant to migration theories in general and forced migration in particular. On top of that, a brief mention of theories related to destination countries of asylum-seekers will be presented. In Section IV the different types of limitations and challenges of such migrant mobility are elaborated along with how Covid-19 outbreak, as a shock event, affected the predictability of the forecasting models. In Section V, the popular forecasting models that have stand out over the years will be reviewed. Extra attention will be put on a specific category; time series forecast models. In Section VI the methodology applied is presented, which is followed by the section of corresponding results. Last, in Section VIII a discussion of the findings and topics for future research will be included and the respective bibliography which provides additional details for the biggest part of the main text.

2 Mixed Migration Background

In this section, mixed migration will be examined as a current phenomenon of research along with its growing importance in Europe and beyond.

2.1 Definition of "Mixed Migration" Flow and the Case of Europe

Irregular migration has become the center of attention for European policy makers. The current crisis' complexity demanded reassessment of the contemporary policies along with an expansion of the available terminology to express its characteristics. "Mixed migration" was proposed to describe this specific migrant mobility. According to the United Nations High Commissioner on Refugees (UNHCR) "mixed movements" or "mixed migration flows" is a complex, hard to predict and monitor migratory population movement including refugees, asylum-seekers, internally displaced, stateless people, economic migrants and other migrants, often moved irregularly (e.g. without the appropriate documentation for their movement) and for numerous reasons. Hence, it is an umbrella term.

The origin behind the intention of asylum applications can be found under three circumstances. It can be an intended purpose of seeking international protection in a given country, it may emerge en route, due to specific circumstances (e.g. access to a country's immigration regulation, political or economic insecurity), or in some cases this decision may arise upon arrival to the destination country.

In the same vein, it is crucial to highlight that "mixed migration" flows exceed the characteristics of a refugee based on the Convention in the Status of Refugees of 1951 where

"a refugee is someone who is unable or unwilling to return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion"

The exact definition of every term that lay underneath the "mixed migration" umbrella term will be briefly described below as defined by the United Nations High Commissioner on Refugees (UNHCR).

Asylum-seekers: people petitioning for international protection, however have yet to be processed. According to Fischel de Andrade (2008, p.114-136), although international protection has been around since the 1917 Revolution in Russia and the exchange of populations between Greece and Turkey, persecution as the reason for protection was vaguely first used in 1938 in the "Convention Concerning the Status of Refugees coming from Germany"[1]. Since then the term has evolved and the mandate[2] of UNHCR has broaden along with it. At the end of 2019, 4.2 million people were asylum-seekers worldwide.

Refugees: people whose asylum claims have been accepted. Refugees before their applications' approval are considered asylum-seekers. 20.4 million refugees were recorded under UNHCR's mandate[2] in 2019.

Stateless people: the international legal definition of a stateless person is *"a person who is not considered as a national by any state under the operation of its law*". People can be born stateless or become one in next stages of their life. Some causes of statelessness are results of gaps in nationality laws and inadequacies in passing down parents' nationality. If laws are not carefully written, some people may be excluded and end up with no nationality. "*Nationality laws in 27 countries worldwide prevent mothers from passing their nationality to their children on an equal basis with fathers. Over 60 countries deny women equal rights with men regarding the ability to acquire, change and retain their nationality, and to confer nationality to non-rational spouses."* (The Global Campaign for Equal Nationality Rights)

[1] "Article 1 of the Convention Concerning the Status of Refugees coming from Germany …The difference … which leaves out from the definition of "refugee", persons who left Germany for reasons of purely personal convenience …" [Fischel de Andrade, 2008].

[2] "The purpose of mandate Refugee Status Determination (RSD) is to permit UNHCR to determine whether asylum seekers fall within the criteria for international refugee protection." (UNCHR, 2011)

Internally displaced: the definition of internally displaced people is the one of people who have not crossed their country's borders to seek for protection but have stayed within their own country even though the government can be the reason for their displacement. 45.7 million people were internally displaced at the end of 2019, according to the Internal Displacement Monitoring Centre (IDMC) due to armed conflict or human rights violations.

Migrants: this term is mostly used to describe people who have moved from a place to another for various reasons. International migrants are individuals who have changed their country of usual residence, whereas internal/domestic migrants change their current location and stay within national states. Another way of characterizing a migrant is based on the different reasons of their mobility, some of which can be economic improvement or education reasons. Lastly, a distinction is made between short-term and long-term migrants based on the duration of the location change. International migrants are not asylum-seekers. Asylum-related migration is an involuntary movement.

The terminology described is approved and commonly used worldwide. However, the European policies and legislation is the geopolitical target of this thesis, hence some clarifications are needed to demarcate the boundaries of the research.

Researchers have approached Europe in different dimensions. Briefly, geographically, Europe consists of 51 independent states with Russia, Kazakhstan, Azerbaijan, Georgia and Turkey located in both Europe and Asia. As a continent of the world, it is part of the United Nations, an international organization founded in the aftermath of the Second World War in 1945 having as the ultimate mission an orchestrated endeavor for international peace and security among other equally crucial purposes. Holy See (Vatican City) and Kosovo are the only European countries that are not members of the United Nations. The Pope's preference to not affect international policy and the fact that Kosovo hasn't gained international recognition by some as an independent country constitute those two reasons. Nevertheless, they all share the same position on policies related to the current migration crisis.

In another dimension, Europe has been mostly associated with the European Union. Initiated as the European Coal and Steel Community in 1950, 6 European countries shaped a union currently known as the European Union and up to 2020 numerates 27 Members States. In order for an independent country to be a member of the European

Union, it needs to be fully or partially located in European territory and meet specific political and economic criteria. European Union established a Common European Asylum System (CEAS) in 1999 and its legislation has been reformed numerous times since adjusting to the needs of every situation. On that ground, European Asylum Support Office (EASO) was founded in 2010 to coordinate the endeavors between the EU Member States on matters related to asylum.

In conclusion, most of the sources mentioned throughout this thesis approach Europe either from a geographical spectrum that involves exclusively counties fully located in European territory hence Turkey, Russia, Kazakhstan etc. are excluded or from a European Union spectrum (e.g EU-28^{+} 28 European Union Members[3] plus Norway and Switzerland) hence some European countries are excluded from their analysis. Lastly, European organizations give great importance to the international aspect of this specific migrant mobility and exclude the internally displaced people among others. Picture 1 shows the position asylum-related migration has in relation to other types of migration. The above fact is considered as one of the most crucial limitations of the mixed migration topic since discrepancies between organizations or nations lead to inaccurate estimations of movements.

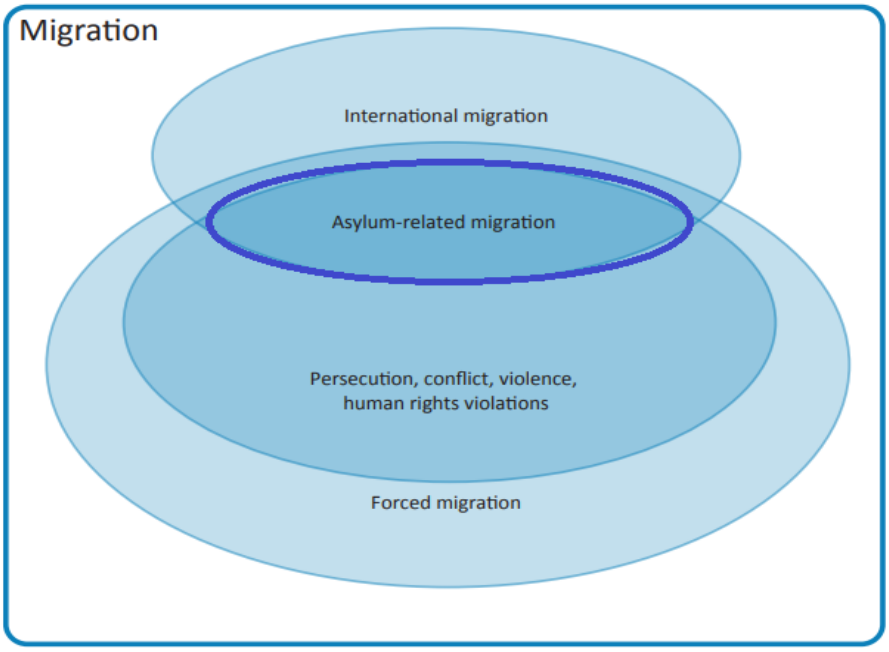


Figure 1: Asylum-Related Migration (EASO,2017)

[3] Since 2020, EU-28 or EU-28^{+} has been replaced by EU-27 and EU-27^{+} since United Kingdom is not a Member State of the European Union.

2.2 Growing Importance

Human movement is a phenomenon as old as the existence of human species itself. It is closely related to the aspect of our behavior. Yet, no adequate handling has been performed to reassure balance between sending and receiving states and societies. Liberty and security are basic human rights. No one shall be subjected to torture and if torture exists, people have the right to seek asylum from persecution since 1948 (UN Universal Declaration of Human Rights, Article 14). However, even though the above fundamental rights apply to all people, world politics has been convulsed by perennial disagreements over forced migration. Current events, such as the Covid-19 pandemic breakdown revealed that humanity is a dynamic organism that is affected holistically and can thrive when it performs as a whole. Humanity cannot thrive if its people die or get abused. Humanity cannot thrive if its people feel not accepted at their place of residence and are separated from family members. Adversities emerge because in this constantly evolving process, sustainability shall also be guaranteed. In other words, the ability to endure over time must be ensured. As Merkel proclaimed in 2018: "*We were always proud of freedom of movement but we never really thought about protecting our external borders. Now we are working on our entry-exit system*". It is a fact that country leaders around the world express the same fears regarding the unpredictable impacts of this new reality. Hence, comprehension in conjunction with efficient policies towards smooth transition and absorption into this new era is crucial considering that this phenomenon seems to expand.

2.3 Particularly in Europe

Undoubtedly, Europe currently faces the largest refugee crisis since the Second World War. According to Gallup's latest update on its Migrant Acceptance Index[4], within the last three years, the global score on the index declined from 5.34 to 5.21 with a number of European countries topping the list of the least-accepting countries in the world. The

[4] "Migrant Acceptance Index is the sum of the points across three question, with a maximum possible score of 9.0 and a minimum possible score of 0.0. The questions are whether people think migrants living in their country, becoming their neighbors and marrying into their families are good things or bad things. 138 countries took part." [Esipova, Pugliese and Ray, 2018].

first Gallup Migrant Acceptance Index was conducted the year following the 2015 migrant crisis in Europe when more than a million people headed to the continent fleeing war and poverty in the Middle East and beyond. UNCHR's reports (UNCHR, 2016) reveal that in 2016 Germany was the largest recipient of new asylum claims followed by Italy and Turkey. In 2017 and 2018 the picture was no different. In 2019, Germany for one more year received the biggest number of new claims in Europe, however the countries that followed were France and Spain. The Migrant Acceptance Index showed that the least-accepting European countries were none of the above. North Macedonia, Hungary, Serbia and Croatia were the least accepting, whereas Iceland ranked on the 2nd position worldwide, followed by Sweden on the 8th position of the most accepting ones. Would it be fair to claim that their level of acceptance is not linked with the number of migrants they receive? The truth is that 85% of worldwide refugees were hosted in developing countries in 2019. This percentage is almost the same since the beginning of the migration crisis in 2015. Reality is, however, that within the duration of the current migration crisis, European reaction towards migration has altered in various ways and this has affected the social harmony. Politicians refer to the current crisis as "humanitarian crisis". Yolanda Onghena (CIBOD, 2015, p. 7) posed two questions in a report about 10 side-effects of the refugee crisis in Europe, concerning numerous rights that have been gradually earned throughout the years such as rights towards minorities and respect for human dignity. These questions are: "*Are these values [respect for human dignity, freedom, democracy, equality etc.] European, universal or national? How does this story of values reach the citizens and to what extend can we consider Europe an engine of motivation?"*. Although, they are multifaceted questions, people need to self-reflect on the subconscious bias they possess before answering those questions.

3 Literature Review

In this section, a review on theories related to international and forced migration will be explored in a chronological order. The understanding of those theories combined with the ones towards destination countries will cultivate a more holistic idea of how those matters have been approached over the years.

3.1 Theories Relevant to International Migration and Forced Migration

Academics and researchers approached different types of migration at a micro- and macro-level. Initial endeavors towards a more thorough understanding of those phenomena, mainly focused on implementation of preexisting theories.

Micro-level approaches have been popular for studying international migration for numerous reasons. Neoclassical economics (Lewis 1954, Ranis and Fei 1961, Todaro 1969, Harris and Todaro 1970, Ahamer et. al 2014) showed a potential theoretical framework with its fundamental principles, however, to be quite unrealistic. The assumptions of the neoclassical theory are that migrants are a homogenous group acting rationally for their financial benefit decision. Sjaastad (1962) approached migration as a form of an investment decision, so individuals based on their individual characteristics (e.g. sex, age etc.) in conjunction with their level of income decide which option is most favorable. When migration is more appealing, it is expected to occur. Although neoclassical theory dominated for a lot of years, recent studies revealed the explanatory contribution of a network theory. A network theory also found in the literature as chain migration (Spinks, 2013) is the case when an individual is influenced by the fact that family members or previous migrants have already chosen a specific country to migrate. Consequently, network theory contributes in a psychological level as well and shows similarities with the herd behavior[5](Banerjee, 1992). Impressively, Docquier et al. (2014) concluded that potential migrants are influenced by the size of the network of previous migrants and the average income in the country of destination. The factor, however, which makes a potential migrant to an actual one, according to the research, is the economic growth of the destination country.

Physical and mental security is taken for granted in the above cases. What happens when violence is present? For the last decades, academics urge the scientific community that international migration theories are not the appropriate tool to explain forced movement because the former is voluntary and the latter is not (Castles, 2003).

[5] Following "popular" choices because other people's decisions may reflect information that they exclusively have.

Nevertheless, before the above finding, some micro theories had been already applied to forced migration which potentially contributed to the realization of a necessary distinction between those two movements. Neoclassical models proved inadequate, because this labor mobility differentiates to the fact that violence increases the costs of staying. A research within the time span of 43 years (1952-1995) showcased that the number of forced displaced people worldwide, has a positive correlation with the level of fear individuals have been exposed to. Consequently, violence outweighed political and economic variables as the dominant drivers of mobility. Some of the asylum-seekers may be poor but this is not the factor that compels them to leave their homes. A network theory is present in this scenario as well, but mostly it influences the country of destination rather than the escape itself. It is important for an asylum-seeker to have family members in a safe country to provide asylum-related information and eventually, if they choose to seek asylum there, reunite with their families. Schmidl (1997) proposed that different forms of expressed violence produce different migration flow results. Applying a pooled time-series analysis from 1971 till 1990, the findings were as follows. "*Generalized violence produce more refugees than human rights violations. On top of that, civil wars without foreign military interventions push fewer people to apply for asylum abroad than do civil wars with foreign military intervention. Lastly, ethnic rebellion does not seem to load to massive population movements, although it is considered a significant cause of refugee migration."* Another factor that can identify the level of emergency of a migration, apart from the form of violence and the existence or not of foreign military, is considered the travel route selected. Economic migrants traverse less dangerous routes than asylum-seekers and are associated with long-term migration (Kalt et. al. 2013).

On the other hand, macro-level approaches attracted a lot of supporters as well. In a macro-level, Piore (1979) a dominant personality in this field, argued that international migration emerges from "*advanced industrial systems due to structural inflation, motivational problems, economic dualism and the demography of labor supply*". Wallerstein (1979), on the other hand, believes that migration is an involuntary reaction to the worldwide expansion of capitalism, the only world system. He proclaims that there is the capitalist West and the Third World South, ostracizing the power of East and its delusional socialist manifesto. Some writers have even drawn a parallel between migration and modern-day slavery. More specifically, in Castles et al. (2014, pp. 32-33) book, international labor migration is dealt as an expression of dominance between counties.

Lastly, an interesting approach can be found in Kick et. al. (2001) where the country-level interplay and its global projection of noneconomic links between countries is presented for two time-intervals.

Papers towards macro-level theory pertinent to forced migration were proved depleted. A worth-reading paper is Castles' (2003), where although forced migration is presented as a result of globalization and is expressed through a North-South inequality, deserves a social framework for analysis. It proclaims that North countries have made refugee-to-be criteria so restricted, allowing no "environmental refugees" or families of internationally displaced from their countries of origin due to development projects. According to Castle, conflict and consequently forced migration will remain present until there is a transformation of both the receiving-sending societies and this phenomenon stops being studied in isolation. Internationally, and most specifically, European countries have proved throughout the years that migration laws that attract explicitly highly-skilled personnel are seen as desirable compared to other groups of labor force. Bansak et. al. (2016) conducted an experiment[6] submitted by voters of 15 European countries with the ultimate goal of categorization of the traits that an asylum-seeker should possess to be approved as refugee by each individual voter. Unsurprisingly, age, religion, host-country language fluency and occupation rank affected the decision of the voters. Although the findings showcased that the asylum-seeker's country of origin did not seem to concern, a more recent examination may produce different results.

3.2 Theories Relevant to Destination Countries

What makes a country appealing to asylum-seekers? Numerous researchers have devoted time in examining the reasons specific countries are preferable than others. Studies revealed that location, smugglers and family networks/historical ties are considered crucial determinants (Spinks, 2012, pp. 2012-12). Smuggling, in particular, has been on the spotlight for numerous consecutive years. Meetings and conversations led to coordinated operations to crack down such phenomena. An example is Turkey's agreement to cooperate with EU and FRONTEX to decrease the smuggling networks (Ruhrmann & FitzGerald, 2017) from the most frequent passage. The emergency conditions individuals are into, makes them rely on smugglers to decide the destination countries or they

[6] The experiment was conducted prior to the current migration crisis.

travel to neighboring countries. For instance, Robinson and Segrott (2002) interviewed 65 asylum seekers in the UK detecting that most of them just prioritized finding a safe place. Kuschminder et al. (2015) presented important papers regarding the determinant role smugglers play in assisting asylum seekers reach a destination country. One of them is a study of irregular Afghan migrants in the Netherlands, over a third of which reported not having chosen their destination because smugglers or relatives had arranged their migration. Hynes (2003) reported a confession from a refugee-to-be regarding smugglers that "*The agent who arranges your flight will also not trust you. He will ask for full payment in advance and he may not even tell you which country he is going to smuggle you into. You will not be told the route, the identity of your guides, or even the identity of your fellow travelers*" (2002b:64). Neighboring countries are also witnessed as potential transit countries due to their accessibility, however are not selected as destination countries. They majority of people are reluctant to believe that the reasons behind this fact are once more related with the asylum-seekers safety and they tend to refer to it as "country shopping" (Spinks, 2013).

A big part of literature expresses the various theories towards receiving and destination countries as "push - pull" forces. This approach assisted academics, researchers and policy makers to acquire a more representative picture of the phenomenon with the ultimate purpose of prediction. The first theory that tried to bridge the "pull - push" forces was introduced by Lee in 1966 where he presented the 4 types of factors that shape mobility. Factors associated with the origin area, factors associated with the destination area, the obstacles en route and personal factors. Nevertheless, opposers of the terminology bring to the table some limitations. Similarly to the neoclassical theory, "push-pull" theories are too simplistic and treat individuals as a homogenous group that reacts similarly to the same stimuli. In addition, peace and violence treated as polar opposites denoted by 0 and 1 simplify the complexity of every situation.

4 Limitations, Challenges and Shock Events

In this section, some commonly observed limitations and challenges will be elaborated. On top of that, out of a real case study, Covid-19 pandemic, the sensitivity of forced migration behavior will be studied in more detail.

4.1 Limitations and Challenges

Forced migration is a global phenomenon that seems to generate more questions than it answers as it is furtherly studied. Setting a framework for analysis and prediction of the most uncertain demographic change (Disney et. al. 2015) can meet a lot of limitations and challenges.

According to Disney et. al. (2015) who conducted a report for the Migration Advisory Committee, the uncertainty around forced migration can be summarized in three components; uncertainty about the future, discrepancies in data and the uncertainty coming from relying on a single forecasting model. Extensively, the first component is observed on the sudden variation in the number of asylum applications after events impossible to predict such as an agreement between countries or a pandemic outbreak. In other words, forced migration is volatile to sock events (wars, recessions, floods etc.) Next, the key ingredient for an accurate prediction is a trustworthy dataset. However errors and discrepancies are witnessed there as well[7]. Last but not least, different forecast models have their advantages and disadvantages. Generally, "mixed migration" flows which consist of both asylum seekers and international migrants, two migrant mobilities that do not share the same stability throughout time cannot be studied with the same forecast model. More specifically, forced migration which appears to be extremely unpredictable, deserves probably combination of methods or even new ones.

Although the above limitations undoubtedly exist, researchers proclaim that are not the cause of this lack of apprehension. Initially, Bansak et. al. (2016) and Ruhrmann & FitzGerald (2017) believe there is a paradox between Europe's endeavor to honor the

[7] Extended information regarding this can be found in the Methodology Section of the paper.

promise towards people of the world in need (Convention Relating to the Status of Refugees, 1951) however, the voters' support should not be affected. The asylum paradox supports the same rationale but it focuses on the efforts European organizations put to help asylum-seekers while keeping them away of European territory at the same time. Prior to refugees' arrival, citizens of Europe and the world made predictions regarding a previous mobility crisis; the migrant workers and their position in, at that time, politico-economic environment. The ultimate goal was cultural assimilation and the policy makers revolved their legislation and policies around this. Reality showed that their policies were unsuccessful once more. Australians who once desired to stay white (Castles 2003, Papastergiadis 2004), became one of the most multicultural society, United States which published the US Immigration Reform and Control Act in 1986 that was designed to decrease illegal migration, led to an increase of both legal and illegal flows (Portes, 1997, p. 818; Castles, 2003) and a European example, Germany, which is currently one of the most diverse societies had as a plan to sent back the migrant workers who were hosted between 1955 and 1973. The truth is that after 11 September 2001 and a set of terrorist attacks, the world has been reluctant to refugees. The truth is also that technological advances have made possible for individuals to have both the cultural identity of their country of origin and the one of the country or residency. Cultural assimilation is again not possible. In conclusion, scientists who study public opinion have revealed that democratic societies follow public opinion remarkably (Burstein, 2003, pp. 29-40.) Forced migration should be seen as a dynamic social transformation that needs to be faced with respect and assertiveness and not as a threat.

4.2 COVID – 19 Outbreak

The infectious disease, COVID-19, escalated rapidly to a global pandemic having impacts not only on an individual level but on a societal as well. European Asylum Support Office (EASO) released an emergency report (EASO, 2020) announcing the new measures and the circumstances under which the process of receiving and evaluating asylum-related applications will be performed. Its rapid spread and the fact that during the winter months additional seasonal diseases are present, made clear that drastic policies should be made. According to this report, 11 EU+ countries declared a state of emergency due to COVID-19. Quarantine and travel ban across countries were some of the measures that were crucial to be taken.

How did those measures affect the asylum-migration flow and perception around this movement in general? EU and international organizations such as EASO, IOM and UNCHR adapted to the new reality immediately, once World Health Organization classified COVID-19 outbreak a pandemic in March 2020[8]. Multilingual information on the epidemiological emergency were created, emergency shelters for either potential coronavirus cases or active cases who needed to self-isolate were available, interviews were undertaken through videoconference and applications were performed online with extended deadlines. Actually, in previous years, within spring months applications tended to increase, however, in 2020 they decreased due to the measures taken to control the spreading of the virus. The percentages of applications in terms of countries of origin were almost the same, with Venezuelans to have a rapid decrease. Syria continued to be the country of origin of 19% of the applications, 9% were Afghans and 8% Iraqi. An approach that altered but seems to be regardless the pandemic is that Syrians' applications were put on top of the list which led to a positive decisions' leap of 22%. The pre-COVID refugee status or subsidiary protection was 30%, and in April 2020 52%. In summary, European organizations took the precautions needed to prevent the spread of the Covid-19 virus and asylum-applications' approval was increased due to the Syrian origin of the applicants.

The pandemic occurred in a part of history that two political events were in force. The first one is the Dublin regulation and the second one is the EU-Turkey agreement. It was observed that asylum-seekers would apply for protection in various European countries causing inconsistent and overestimated records of forced migration. A solution to the above challenge was the implementation of the European Commission's EURODAC and the Dublin regulation. EURODAC regulation concerns the unique identification of applicants based on which, an EU asylum fingerprint database was supposed to be created to store, above the age of 14, applicants' fingerprint and cross check when a newly-arrived applicant case appeared. In addition, the Dublin regulation put on the table the fact that only one EU member state should be eligible to examine an asylum-related case based on 3 criteria. Those criteria are in hierarchical order, family considerations, recent possession of visa or residence permit in a EU member state and whether the entry to the country is regular or irregular. The EU-Turkey

[8] WHO announces COVID-19 outbreak a pandemic. Retrieved from https://www.euro.who.int/en/home

agreement in March 2016 between the European Council and Turkey, agreed to stop the flow of irregular migration via Turkey. Turkey agreed to receive back any irregular migrant crossing from Turkey in exchange to lifting the visa requirements for Turkish citizens, additional funding to Turkey and a Voluntary Humanitarian Admission Scheme to be activated when stabilization of the condition is ensured. Due to the pandemic, some borders were closed among EU members. Although the Dublin regulation is still in force, the European Commission's Guidance published that *"where a transfer to the responsible Member State is not carried out within the applicable time limit, responsibility shifts to the Member State that requested the transfer pursuant to Article 29(2) of the Dublin III Regulation. No provision of the regulation allows to derogate from this rule in a situation such as the one resulting from the COVID-19 pandemic"*. Additionally, in late February, Turkey attempted to open the border with Greece.

In a more general effect, the pandemic had a behavioral impact towards asylum-related migration. Research around this observation was collected and presented in a report (EASO, 2020) where the following findings are mentioned. The evolutionary theory supports that individuals are skeptical and concerned towards strangers, who may be carrying developed infections. As a result, xenophobia is a phenomenon witnessed in a greater extend during disease outbreaks. Lastly, Beall at al. analyzed voting intentions before and after the Ebola outbreak and observed that the after-Ebola voting intention was increased for a conservative candidate. This explain, to some extent, why liberal countries are more accepting towards asylum-seekers compared to Western Balkans countries and Past-Soviet States.

5 Recent Theories Relevant to Forecast

In this section, the deterministic and probabilistic approaches towards forecasting will be showcased. Time series forecasting methods is the collection of models in interest in this thesis.

5.1 Deterministic Approaches

Migration flows record has been a way for policy makers, media and citizens to have access to the level and direction of population movements. The free movement of European Union citizens had resulted in the publication of Community Regulation No 311/76 which required members of the European Union to supply migration statistics annually to Eurostat since 1976. Those records service another pervasive need; forecast. Worldwide attention has been put on this subject, thus, domestic and international projects seek actively to create a useful framework taking into account the limitations and challenges that exist. However, as Henri Poincare once proclaimed, "It is better to foresee even without certainty than not to foresee at all".

In the end of the 20th century, researchers such as Chojnicki (1977) believed that in socio-economic forecasting, indicated as nomological, accurate predictions were guaranteed when the phenomena under study relied directly on theories or laws. This implies that the theories should be inclusive and robust. Öberg and Wils (1992, pp. 6–7) noticed, however, that regardless of the theories at hand on migration, the predictability of this mobility was profoundly limited. 10 years later, in 2002, Kupiszewski expressed the same concerns. The core limitation is that available theories are too simplistic and incapable to explain complex situations such as international migration let alone forced movement. This combined with the fact that some theories involved complicated mathematical operations, allowed fewer forecast models to be explored and, potentially, become a good fit.

Literature regarding deterministic approaches will be cited briefly, since their impracticality serves no further research purpose. In 2006, Bijak collected and presented five versions of deterministic models. The judgmental scenarios where the plausible development of the variables in question are based on a specific theory. The Delphi method and surveys among experts where the favorable theory in every situation is chosen by the forecast experts themselves or by the use of a Delphi method, a way to satisfy group communication. The fact that until 1974 the number of studies on the Delphi method reached four digits (Mahajan, 1976) showcased the academics' eagerness around this method's potentials. Nevertheless, subjectivity of individuals and unquantified terms such as "good respondent group" resulted in both positive and negative experiences throughout the years. The third one is through assessment surveys. Mostly, the findings of assessment surveys are not presented as forecasts. Apart from this, available

papers (Prus j & Johnson R, 1994) warn regarding the difficulty behind constructing good surveys or questionnaires. Then, the mathematical models of migration exist which stem from demography and human geography. Usually the classical demographic models use multi-dimensional linear algebra and calculus in a deterministic way, whereas geographic models are mostly based on the Markov chains which is actually a stochastic approach. Most of these versions has been performed in the 1990s. Although, a lot of study was around deterministic models in 1990s, Keilman (2007) confessed that unfortunately up until 2007 nearly all official forecasts were deterministic, with Statistics Netherlands to be the only known exception. His work on the population projection of the United Kingdom showed that the available approaches underpredicted the international migration to the United Kingdom. Taken for granted the existence of errors, the same report showed that the errors in countries of Northern Europe, were smaller than the ones of Central and Southern Europe, such as Germany, Austria, Portugal and Switzerland.

5.2 Why population forecasts should go probabilistic?

Invention is not a process based on parthenogenesis. Breakthroughs in statistical analysis are no exception. Although probabilistic approaches, in most cases, were derived from deterministic measurement concepts, they managed to move a step ahead on the accuracy, efficiency and inclusiveness. The best-case scenario would be a holistic model as the supportive tool for every decision-making case. Unfortunately, until this day this has not been achieved. Why? Because of uncertainty. When the outcome is observed, a deterministic approach is a suitable approach. In other words, all possible deterministic scenarios constitute a probabilistic scenario. Hence, a probabilistic model includes two elements; a range of different outcomes and probabilities attached to them.

Forced migration is so irregular that numerous studies have failed at making accurate forecasts using probabilistic models let alone deterministic ones. One thing is for sure, forecast and uncertainty go hand-in-hand. Keilman et. al (2002) revealed the initial steps towards probabilistic modelling rationale. In the early 2000s Statistics Norway would have to deal with cases where uncertainty was expressed unquantified. For instance, the number of children in the ages of 6-12 in Norway in 2010 was estimated to be between a number interval depending on the level of fertility. Fertility could be either

low or high. However, the fact that no probability was assigned to any of those two scenarios subtracted flexibility from the government. Knowing that high levels of fertility is 90% possible to occur would have put the policy makers in a position to occupy more teachers, provide citizens with after school facilities etc. and, in general, manage supplies effectively. For the above and more reasons, migration should be approached by probabilistic models.

5.3 Probabilistic Approaches

In this part, the most widely-used probabilistic forecast models and their case study applications will be presented. Disney et. al. (2015) divide those models into 5 categories; time series, expert-based, Bayesian, econometric modeIs with covariate information and historical errors models. Furthermore, after testing all those 5 methods on different migration data, the conclusion was that only two of them are applicable on asylum-related datasets; extrapolation of time series and Bayesian expert-based models.

5.3.1 Econometric Models

Econometric models are widely used in Economics. Econometrics means economic measurement; applying statistical techniques to relevant data with the main purpose to reveal the relationships among economic variables (Kacapyr, 1996). For forecast purposes, it was greatly used to estimate the migration to Western Europe following the enlargement of the European Union (Disney et. al., 2015). An interesting case was Chi & Voss (2005), where a hierarchical regression approach was presented. Zudelova, T.Z.M (2015) compared two econometric models to predict migration flow among the Slovak districts. Despite their popularity, econometric models face some limitations that diminish their predictability power. The model created cannot be applied to different samples. In the case of migration, in a different country. In addition, it is nearly impossible to identify the correct explanatory variables and almost all econometric model constructed is miss-specified.

5.3.2 Expert-based Models

Expert-based models are an appropriate collection of probabilistic models when shock events are present. Switzerland's tool for migration flows' prediction has been expert input for more than a decade (MPD, 2018). Its competitive advantage is mainly that experts have access to information on current and past political, economic and social con-

ditions and additionally their professional experience gives them the ability to foresee changes of events. The fact that data are not of any use in this scenario is a major limitation. This is the reason why, lately, expert-based opinions are combined with statistical models such as Bayesian models. In these cases, even if experts are not able to predict a structural break such as a war, they can adjust the statistical model to increase its future predictability. An additional challenge can be the presence of bias, hence there are cases where group of experts of different academic backgrounds are selected to provide their knowledge input (Weible, 2008, pp. 615-635).

5.3.3 Bayesian Models

Bayesian combined with expert-based models constitute one of the two appropriate class of probabilistic models that has shown promising results on asylum-related migration. Papers on cases of a solely Bayesian modelling application were scarce (Azose & Raftery, 2013; Abel et. al, 2013). Both methods though, use the Bayesian Theory[9] and Bayesian Law. Bayesian approaches allow different parameters to be included and solve delicately the overfitting problem because they allow certain priors on the variables. This is where expert opinions can be included. They can constitute the prior distribution of different parameters. When this happens, the model is transformed from a Bayesian model to a Bayesian expert-based one. An interesting Bayesian expert-based case study is Bijak and Wisniowski's work (2010) on forecasting immigration for seven European countries until 2025. The expert-based contribution of the endeavor was the results of a two-round Delphi survey of migration experts.

5.3.4 Time Series Methods

Time series methods are in the center of this thesis, hence a more detailed and holistic approach will be followed. Time series is a sequence of numerical data points in successive order. The independent variable is the time and the main purpose is to use the dataset for prediction. It is solely relied on past data and the fact that data are imperfect or

[9] $F(Y_t|\beta,\sigma^2) = (2\pi\sigma^2)^{\frac{-T}{2}} \exp\left(-\frac{(Y_t-\beta X_t)^T(Y_t-\beta X_t)}{2\sigma^2}\right)$,

Retrieved from https://towardsdatascience.com

incomplete introduce bias and errors into the forecasts. It is crucial for the model to identify a possible trend[10] and seasonality[11], however, shocks or agreements, such as an economic crisis cannot be added which potentially may result in a false prediction. In previous part of the paper, it was mentioned that time series extrapolations can be an applicable model for asylum-related migration, however similar models exist which do not fall under the same category (e.g. time series extrapolations through propagation of historical forecast errors). The above version can be applied to other forms of migration prediction. Asylum-related migration is almost always a non-stationary process which requires special handling in regards to forecast time horizon. In this paper (IOM, 2020) general migration forecast horizon is proposed to be up to 100 years, however in cases such as asylum-related migration it is urged that the time horizon should be up to 1 year. Surprisingly, in Carammia & Dumont's (2018) paper the medium to high predictability is correlated with a forecast horizon of a week or a month. Prediction a year ahead is correlated with low to medium predictability.

The special characteristic of time series analysis is the fact that observations, consecutive ones or not, are not independent and if their pattern is studied, inferences can be drawn from such series. Acknowledging the dynamic and static forces of a system along with their association, assists in comprehending the system's mechanism and, eventually, estimate its next phase with some level of certainty. This uncertainty is observed in stochastic time series where the future values have a probability distribution which is conditioned by a knowledge of past values e.g. $x_t = 4x_{t-1} + \varepsilon, \varepsilon \sim N(0, \sigma^2)$

The definition of a time series model for the observed data {xt} is a specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables {Xt} of which {xt} is postulated to be a realization (Brockwell and Davis, 2016).

The various forces mentioned are called components of a time series and are three; trend, seasonality and randomness. A system can exhibit one, two or all of them. A time series that possesses all of the three forces can be witnessed in Figure 2.

[10] More about trend analysis can be found in the Appendix.

[11] More about seasonality analysis can be found in the Appendix.

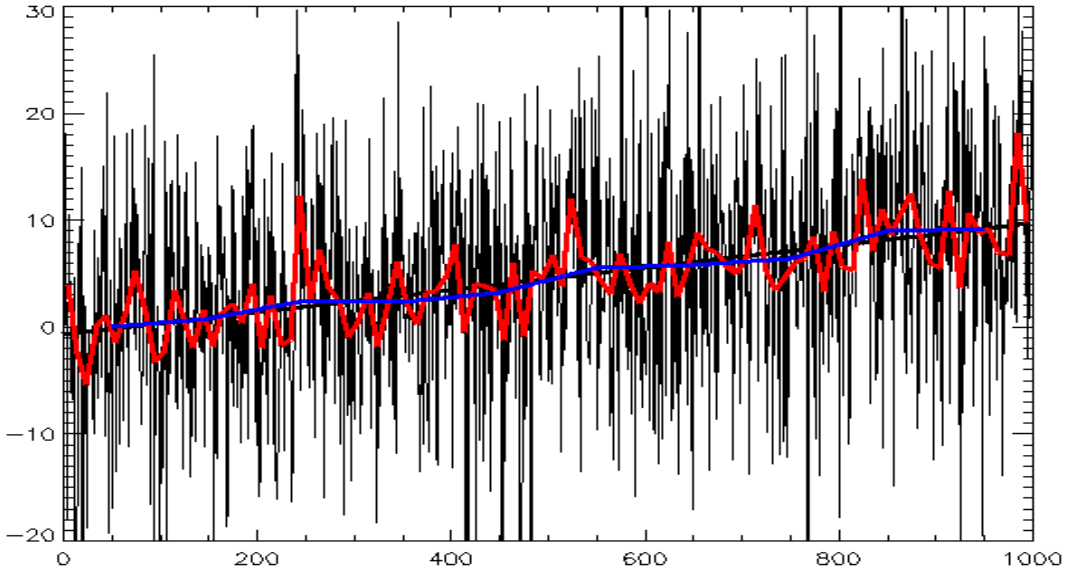


Figure 2: Time series graph[12] (red line: seasonality, blue line: trend, black line: randomness)

6 Methodology

A study research needs a dataset and a rationale behind the thought process. Appropriate data is one of the most important ingredient for an accurate forecast. National, European and international agencies track almost every quantitative part of everyday life for data-driven decisions. Europe benefits from internal statistics provided by EUROSTAT, EASO etc. and external ones such as UNHCR. Nevertheless, challenges are present throughout the process. Inadequate resources and monitoring equipment combined with inconsistent terminology over the years and between countries can potentially produce untrustworthy datasets. In this section, the challenges of selecting the most appropriate dataset will be examined along with the methodology used to collect the specific ones on a country-wise level and source-wise level.

Asylum-related migration belongs to a distinguished category because it is a highly political topic and the international component is clear. Issues emerge on an individual, institutional and even societal perspective. Bakewell (1999) mentions that over the years the United Nations High Commissioner for Refugees (UNHCR) and the US Committee for Refugees (USCR) estimated different numbers of refugees and asylum

[12] Retrieved from https://en.wikipedia.org

seekers. The author claims that even if an agreed terminology is ensured, disagreements will emerge on whether individuals meet the criteria. More than 10 years later, the same adversity still existed. In Mouzourakis' report (2014), the author reveals that EASO, EUROSTAT and FRONTEX released different numbers of asylum applications for 2013. A UNCHR report (Hynes, 2003) on issues of mistrust between authorities and refugees in the UK, argues that obtaining data and defining a "real" refugee by authoritative sources is not optimal. According to the report NASS[13] operates believing that the majority of asylum seekers do not worth the refugee status. Another report (Ljones, 2019) reveals that "*Refugees and refugee statistics will be phenomena that are closely linked to severe conflicts – even wars. This … that statistics is produced in a professional way and that political inference has no effect on the statistics*." Additionally, it is highlighted the importance of both trust in the quality of statistics and the principles of confidentiality towards the asylum-seekers. Overall, factors to be considered are bias towards to-be refugees, terminology inconsistency and political motives.

Finding the most appropriate tools for a database assessment has been a trial-and-error process and it is still developing. Methods that are used today, might be obsolete in a couple of years forward. Prior to the EURODAC database, asylum-seekers whose application was rejected, would re-apply elsewhere, leading to double counting in asylum statistics. As Mouzourakis (2014) revealed based on an EASO's report "Although the proportion of outgoing requests was on average about 12% of the number of registered asylum applicants, Dublin transfers were made in the case of only about 3% of those making an asylum claim in the EU", leading to further discrepancies.

Interestingly, EASO published a methodology for evaluating potential data sources (EASO, 2017). The evaluation was based on 6 criteria; frequency, definition, coverage, accuracy, timeliness and quality assurance processes. A database that would meet those criteria perfectly, would publish real-time statistics reflecting to the fullest the definitions agreed on asylum-related migration with no errors. The reality is that a perfect database does not exist. However, in 2007 the European Parliament and Council published the Regulation (EC) No 862/2007 which adopted harmonized rules "concerning the compilation of European statistics related to migration and international protection (asylum)" (EASO, 2011).The asylum-related statistics criteria would be collected based on

[13] NASS stands for The UK National Asylum Support Service.

geography, year, timeliness and data providers. The geographical coverage consisted the 31 EU/EFTA[14] countries, the start year of harmonized data collection was 2008, the frequency and update period would be 2 months after reference period for monthly and quarterly data and 3 months after for annual data. Thus, the frequency would be monthly, quarterly and annually. Lastly, the data providers would be each countries' Ministry of Interior or related Immigration Agencies. The asylum applications of 3 agencies will be explored to showcase any inconsistencies between the two EU agencies after the Regulation No 862/2007 and the international one. EASO, EUROSTAT and UNCHR publish annual data on asylum-related migration and are ranked 1st, 2nd and 3rd respectively on the evaluation matrix based on the 6 criteria; frequency, definition, coverage, accuracy, timeliness and quality assurance processes. The total asylum-related applications were recorded for the year 2011 to explore their characteristics. An important reminder is that UNCHR is an international organization, whereas EASO and EUROSTAT two European ones. EUROSTAT has as a starting point the year 2008, after the harmonization hence an indication of a point before 2008 was not feasible. Based on the above references, the results are expected and shown in Table 1.

It is witnessed that in both cases the UNHCR agency recorded fewer applications. The percentage of difference interval[15] between EUROSTAT and UNHCR seems to be consistent throughout the years, hence there is a portion of movement that is not recorded by UNHCR.

Table 1: Monitor of asylum seekers by UNCHR, EASO and EUROSTAT

	UNCHR	**EASO**	**EUROSTAT**
		2011	
GERMANY	45.717	53.300	53.240
		2017	
ITALY	126.457	128.850	128.850

[14] EFTA stands for The European Free Trade Association. In 2007 they were Iceland, Liechtenstein, Norway and Switzerland. (Retrieved from http://efta.int)

[15] More information in the Appendix.

The author of this paper chose EUROSTAT as the preferred source among EASO, UNCHR and IOM and National Agencies. EUROSTAT publishes monthly data on asylum-related migration which assists in increasing the sample size. The final sample size was 153 observations, from January of 2008 till September of 2020. In a country-wise level the countries that were selected are, namely: Germany (GER), Italy (ITL) and Norway (NO). The countries were not selected based on a specific criterium, however they were not selected completely randomly as well. The main rationale was to find countries located in the north, central and south part of Europe, reflecting on different climate conditions, distance between the major sending countries (Figure 3). Next, two central and south counties were selected since Eastern Mediterranean routes is the most frequently used route toward Europe and a severe victim of the COVID-19 pandemic. Lastly, since the EUROSTAT database was determined, the countries should be included in the 31 EU/EFTA countries in order to take full advantage of the database for the purposes of the paper.

At this point, some remarks will be made for the data downloaded. The original data refer to monthly asylum and first-time asylum applicants by citizenship, age and sex since January 2008 till September 2020 (153 observations). The sex profile consists of two genders (male and female) and the country of citizenship of the applicant is not restricted in any way. All the recognized countries, European or not, are included. Lastly, the topic of interest is new incomes of asylum-related applications. The overall number of applications depend on new and pending cases. Pending cases are affected by many variables such as bureaucracy issues, available staff and other extrinsic factors, however it is a number easily measured and manageable. Additionally, all three of them need to registrate the application for international protection within 3 days after the asylum-seeker entering the European territory, so the new cases per month represent the actual numbers of each month. Hence, only first-time applicants are taken in consideration. The measure of prediction accuracy was selected to the be Mean Absolute Percentage Error (MAPE) due to its popularity and simplicity. An upper limit has also been set to 45% MAPE in order for the models with the greatest performance to be mentioned and studied.

Figure 3 Location of NO, GER, ITL

6.1 Exponential Smoothing Methods

Exponential Smoothing Methods were first proposed in the late 1950s (Brown, 1959; Holt, 1957; Winters 1960) and are a collection of models that are used for forecasting purposes. Their main principle is that recent values hold more predictive power and importance than past ones. Exponential Smoothing Methods assign exponentially decreasing weights as the observations get older. They combine Error/Residuals, Trend and Seasonal components in a smoothing calculation. Those components, when they exist, can be combined either additively or multiplicatively. In R programming language those models can be of use with the forecast package. The general form of every possible exponential smoothing model is of the form ETS (X, Y, Z) with E corresponding to the error, T to trend and S to seasonality. The variables inside the brackets can take either the letters A for additive, M for multiplicative or N for not included and they show the way components influence the time series. Hence, some of the possible models can be ETS (M, N, N) which correspond to Simple Exponential Smoothing (SES) Method with multiplicative errors. This would be the simplest case where no trend or seasonality seems to be present[16] (Figure 4). When a trend is present, a possible model is ETS (A, A, N) which is translated

[16] Formula $\hat{y}_{t+h|t} = ay_t + a(1-a)y_{t-1} + \cdots, for\ h = 1,2,\ldots$

to Holt's Linear Method with additive errors[17] (Figure 5). Lastly, a seasonal time series can be of the form ETS (A, N, A) which is a Holt-Winters seasonal method with additive errors[18].

As mentioned before, Exponential Smoothing Methods combine the three components in a smoothing calculation. The smoothing parameters are three (a, β, γ) and their existence depend on how many of the time series components are present in the specific time series. In the simplest case with no trend and seasonal component there is only one smoothing parameter (a), however when all the components are present the smoothing parameters are all of them. The smoothing parameter in each case is between 0 and 1 and the closer the value of the parameter is to 1 the heavily the time series relies on recent data.

In Exponential Smoothing Models the criteria that determine which model is more appropriate are Akaike's Information Criterion (AIC), AIC corrected for small sample bias and Bayesian Information Criterion (BIC). The forecast package has selected Akaike's Information Criterion (AIC) as the dominant one[19].

---

[17]Formua

$$\hat{y}_{t+h|t} = l_t + hb_t,\ \ l_t = ay_t + (1-a)(l_{t-1} + b_{t-1}),\ \ b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$$

[18] Formula:

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)},\ \ l_t = a(y_t - s_{t-m}) + (1-a)(l_{t-1} + b_{t-1}),\ b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1},\ s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m},\ \ k = (h-1)/m]$$

[19]Formula

$$AIC = -2\log(L) + 2k$$

$$L = likelihood\ of\ the\ model, k = \text{the total number of parameters and inital states that have been estimated})$$

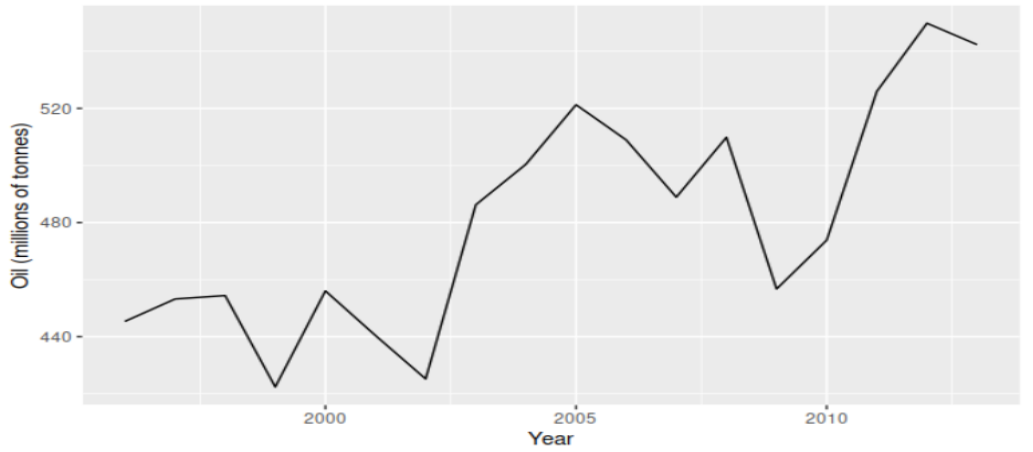


Figure 4: Oil production in Saudi Arabia from 1996 to 2013. No trend, no seasonality.[20]

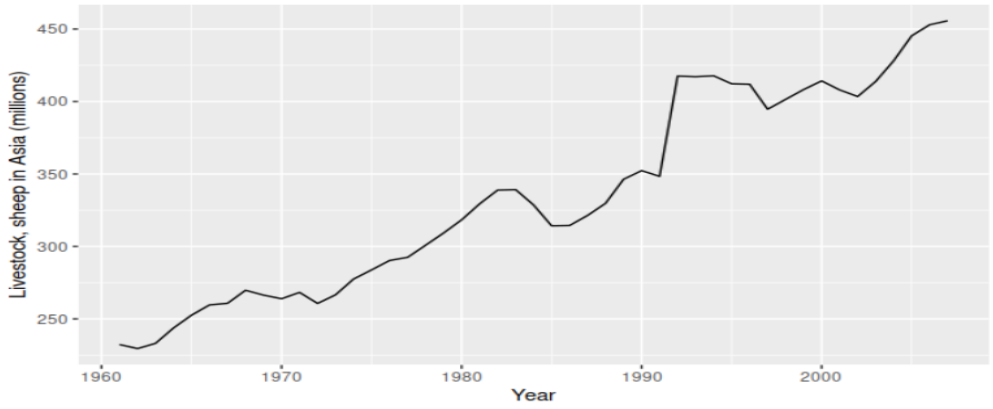


Figure 5: Annual sheep livestock numbers in Asia (in millions). Trend is present[21].

6.2 (S)ARIMA

ARIMA stands for “Auto-Regressive Integrated Moving Average" models and are considered the most general class of models for forecasting non “stationary”[22] time series. It is a class of models that support that previous observations can alone be used to predict the future values. When the time series is stationary, the “Integrated” term is not required and the model is transformed to an ARMA model. ARMA(p,q) models are a

[20] Hyndman & Athanasopoulos, 2018

[21] Hyndman & Athanasopoulos, 2018

[22] More information on stationarity in Appendix.

linear combination of MA(p)[23] and AR(q)[24] models. The general forecasting equation is of the form

$$\hat{y}_t = \mu + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q},\ e_t \text{ is white noise}$$

In other words, the general formula is a constant and/or a weighted sum of one or more recent values and/or a weighted sum of one or more recent values of the error. If the above formula is used for a non-stationary time series there would be a slight difference; the observations would need to be differenced[25] first. This step is important because the term "Auto-Regressive" means that it is a linear regression model and the past observations should not be correlated. However, over-differencing will negatively affect the model parameters as well. In R programming language the package forecast will be once more the tool to test whether an ARIMA model is a good fit for the time series. Generally, the form is ARIMA (p, d, q) where p corresponds to the order of the auto-regressive part, d to the degree of first differencing and q to the order of the moving average part. When seasonality is present ARIMA can model its influence as well. A seasonal ARIMA model is constructed by adding some terms. ARIMA (p, d, q) (P, D, Q) [M], where P is the number of seasonal autoregressive terms, D the number of seasonal differences, Q the number of seasonal moving average terms and M the number of observations per year.

6.3 Neural Network Methods

Artificial Neural Networks were first proposed in 1944 by Warren McCullough and Walter Pitts and consist of thousands of units that are interconnected, trying to resemblance the function of a human brain. Those units are arranged in series of layers, each of which connects to the layers on either side. In the simplest version, the layers are three and are called the input layer, the hidden layer and the output layer. The importance between one unit of a layer and another unit of the next or previous layers is represented by a number called a weight.

---

[23] More information in the Appendix.

[24] More information in the Appendix.

[25] The formula of d^{th}differentiation is $y_t = (Y_t - Y_{t-1}) - (Y_t - Y_{t-2}) - \cdots (Y_t - Y_{t-d})$ (Nau, 2017)

In R programming language neural networks can be found in the forecast package and the appropriate model is represented by the form NNAR (p, P, k) where NNAR is the abbreviation of Neural Network Auto-Regression Model. The variables in brackets correspond to the number of lagged values that are used as inputs, the number of seasonal lags and the number of hidden units respectively. The method's fundamental limitation is the need for large datasets to prevent it from overfitting. In cases where frequent fluctuations are present, the model will fail in generalization.

6.4 Bayesian Methods

Bayesian Structural Time Series model is a statistical technique for time series prediction and is applied to numerous scientific fields. In R programming language, it is available using the bsts package. It avoids delicately overfitting and it has proven to be a effective choice in short or average datasets.

7 Results

In this section the results of the study will be demonstrated.

7.1 Germany

Germany is the European country which received in absolute numbers the biggest majority of asylum claims since the beginning of the current crisis. The number of *annual* first-time asylum applicants since 2008 is displayed below (Figure 6). In the beginning of 2014, the number of claims appear to increase with its peak being around 2016. The main reason lies behind the closing of the Balkan and Mediterranean routes that asylum-seekers used to get to Europe.

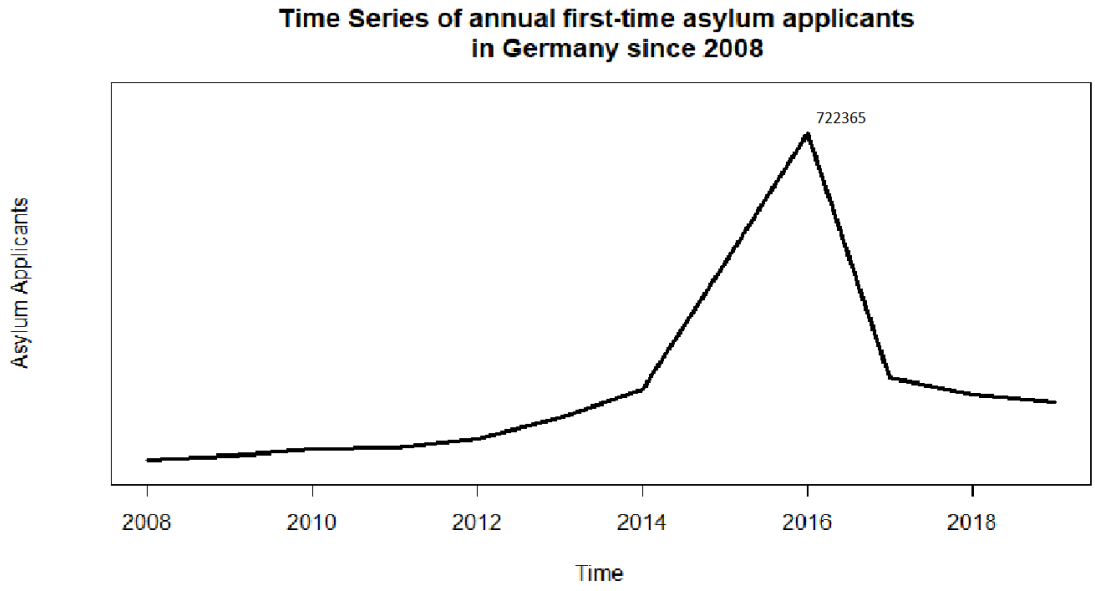


Figure 6 Annual first-time asylum applicants in GER

The Exponential Smoothing Method that generalized better in the monthly time series forecast is the Holt method with the majority of the alternative methods having a small deviation.

As observed below, Simple Exponential Method held the last observation of the train dataset, applied a smoothing parameter and kept it unchanged throughout the time horizon of 9 months. This function of the model can be proved efficient in times of relative stability. However, if the last observation is a product of anomaly such as an economic recession, the prediction will be subject to systematic error. Consequently, its major limitation lies in the fact that the selected time horizon is not taken into account. Especially, in a phenomenon linked with frequent fluctuations such as forced migration, such a method is soon to be found obsolete and replaced with more complex models. In the case of Germany, the last observation turned out to be a fortunate pick and produced a Mean Absolute Percentage Error below the chosen threshold.

In the table below (Table 2), the actual number of asylum claims per month is presented as is the predicted numbers from the Simple Exponential Method and the subsequent error between the two. Following is the time series graph representing monthly first asylum claims in Germany and its contradistinction with the predicted values is shown.

Table 2: Results of SES in GER

	Actual Numbers	**Predicted Numbers**[26]	**Error**
January	13385	8360.28	0.37
February	11140	8360.28	0.25
March	7935	8360.28	0.05
April	5645	8360.28	0.48
May	3775	8360.28	1.21
June	4790	8360.28	0.74
July	7590	8360.28	0.10
August	7275	8360.28	0.15
September	9300	8360.28	0.10
MAPE			**38%**

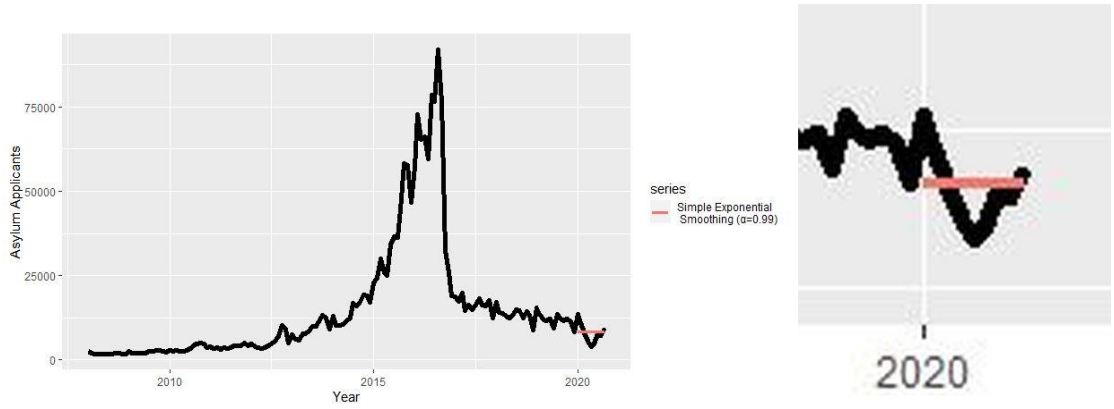


Figure 7 Graph of SES in GER

The Holt Exponential Method although admittedly a more sophisticated model than the Simple Exponential one is still a long way from presenting real complexity. A trending time series can be deciphered by such models. Nevertheless, in this specific time series an apparent upward or downward motion is not present hence its effect is almost unnoticeable on the predicted values. Undoubtedly, the number of asylum claims de-

[26] Last observation is 8360.

creased in 2020 due to the Coronavirus outbreak and so the minor downward trend that Holt method took into account led to a better predictive performance.

In the same vein as above, the table containing the actual and the predicted values along with their respective errors is showcased below (Table 3). Afterwards, the graph consisting of the actual values and the predicted ones is also presented (Figure 8). It is true that the effect of the estimated downward trend is barely noticeable in the graphical representation however one has to look no further than the numerical values of the table to witness that a driving force is active.

Table 3 Results of Holt in GER

	Actual Numbers	**Predicted Numbers**	**Error**
January	13385	8313.68	0.37
February	11140	8276.41	0.25
March	7935	8246.60	0.03
April	5645	8222.73	0.45
May	3775	8203.65	1.17
June	4790	8188.38	0.70
July	7590	8176.16	0.07
August	7275	8166.40	0.12
September	9300	8158.58	0.12
MAPE			**37%**

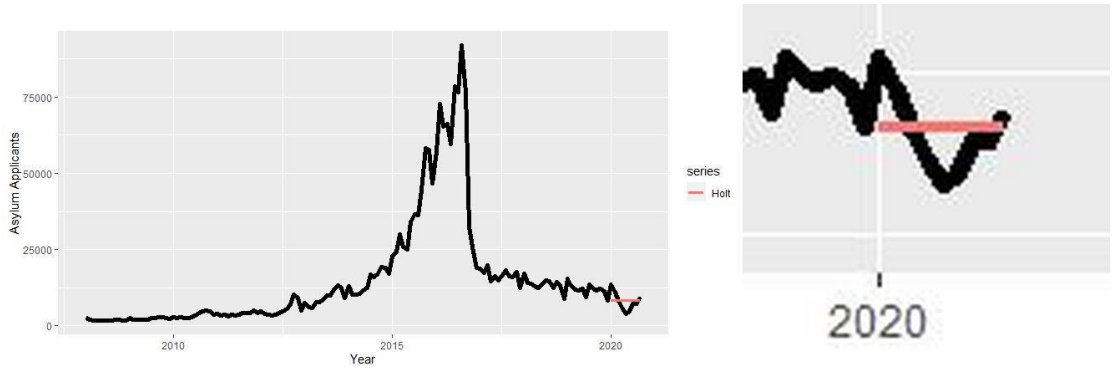


Figure 8 Graph of Holt method in GER

The last Exponential Smoothing Model that achieved a Mean Absolute Percentage Error of less than 45% is the one generated by the automated option given by the forecast package in R programming language. The automated model is of the form ETS (M, N, M) which corresponds to a model with no trend and multiplicative errors and seasonality. Surprisingly, the overall error is higher than the one of the previous models. This is due to the fact that the selection of the optimal model is made by the comparison of the various types of errors (ME, RMSE, MAPE etc.) on the training set which is not a good criterion for a model with good generalization capabilities. The training MAPE of the automated model is 10% whereas the training error of the Holt model is 27% and is thus disqualified. The apparent success of the automated model, though, stopped at the training set as it was proved to be inadequate in predicting realistic values.

The respective table (Table 4) and graph (Figure 9) representing the above findings are the following. Lastly, a graph (Figure 10) containing all the aforementioned models along with the Holt-Winter one is presented below. The Holt-Winter method produced a MAPE of 100% hence its only used as a comparison tool and not as a potential model of choice.

Table 4 Results of Automated in GER

	Actual Numbers	**Predicted Numbers**	**Error**
January	13385	10498	0.21
February	11140	8961	0.19
March	7935	8679	0.09
April	5645	8246	0.46
May	3775	8062	1.13
June	4790	8811	0.83
July	7590	10016	0.31
August	7275	10637	0.46
September	9300	11047	0.18
MAPE			**43 %**

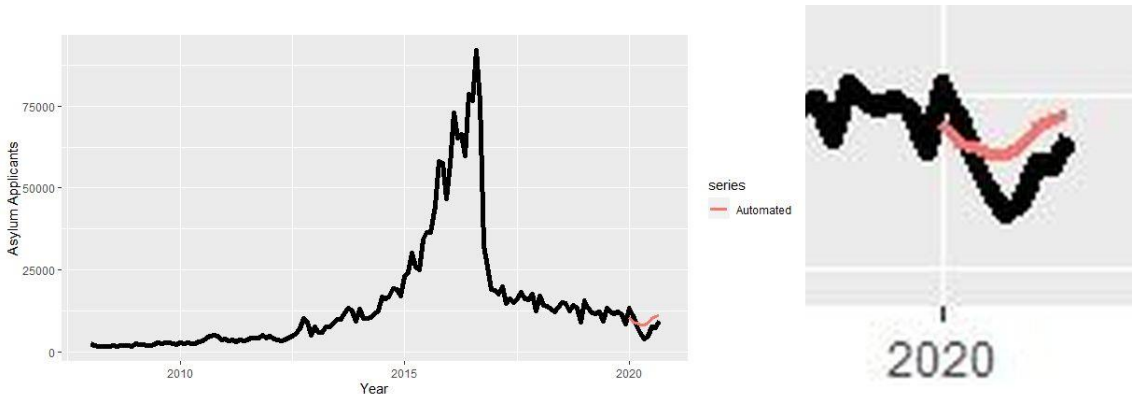


Figure 9 Graph of Automated model in GER

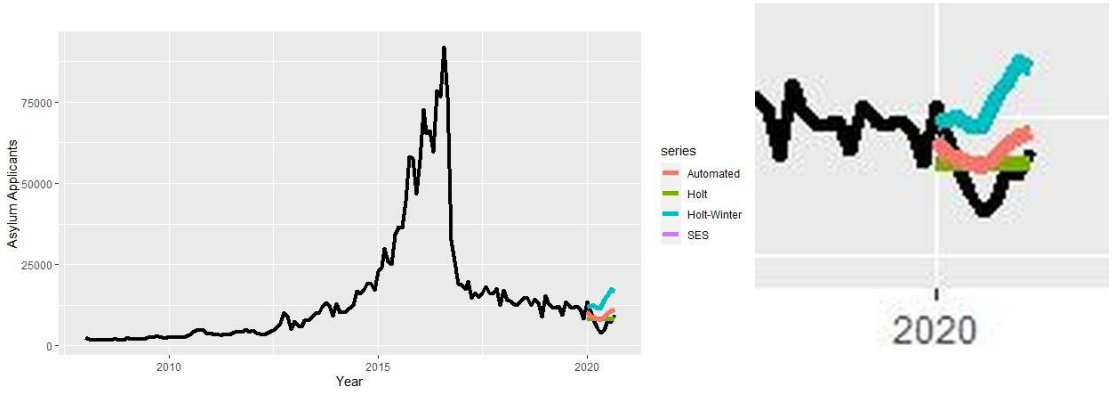


Figure 10 Graph of all Exponential Smoothing Models in GER

At this point it is important to note that it is desired for a model to follow closely the time series' patterns, however generalization must also be ensured. The graphs (Figure 11) below show each model's endeavor to imitate the behavior of the specific time series.

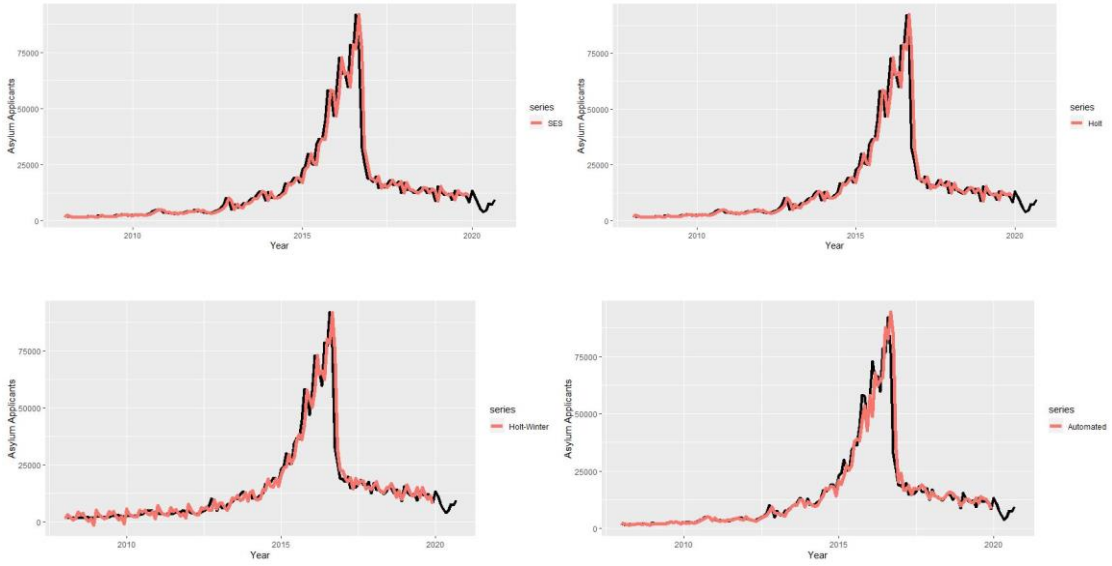


Figure 11 Graphs of all models-fitted in GER

Next class of models in the list is the ARIMA models. As every component of an ARIMA Model is explained before, in this case the appropriate one was proven to be of the form ARIMA (0,1,0). The model used the last observation of the training set to predict every future value. The challenges of this specific model which do not apply to all ARIMA models are the same as the ones of the Simple Exponential Smoothing model.

The table (Table 5) and the graph (Figure 12) concerning the ARIMA model are presented below.

Table 5 Results of ARIMA(0,1,0) in GER

	Actual Numbers	Predicted Numbers	Error
January	13385	8360	0.37
February	11140	8360	0.25
March	7935	8360	0.05
April	5645	8360	0.48
May	3775	8360	1.21
June	4790	8360	0.74
July	7590	8360	0.10
August	7275	8360	0.14
September	9300	8360	0.10
	MAPE		**38 %**

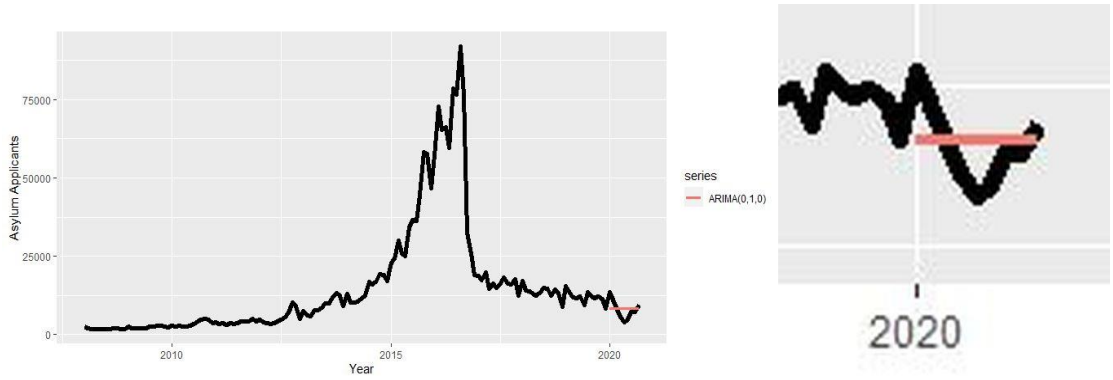


Figure 12 Graph of ARIMA in GER

Neural Networks produced a nontypical result in regards to their usual predictive performance. Their potential to be a good tool in complex cases in various fields, justifies the growing dominance in the machine learning field. This is not the case in this analysis. Neural Networks' MAPE was 69% which was probably a result of overfitting. A fundamental ingredient for the optimal function of a Neural Network is the possession of a large dataset. The training dataset consisted of 144 observations which is considered relatively small. This observation may not be as peculiar as it seems since concerns regarding the compatibility of neural networks and time series have been reported

in the past. Nonetheless, a graph (Figure 13) has been plotted to make visible the challenges that this class of predictive methods face.

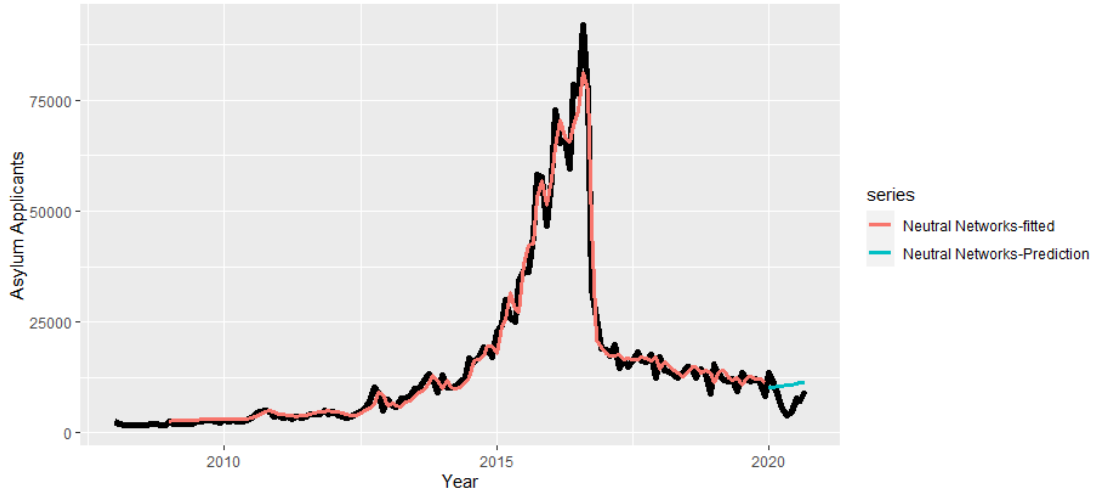


Figure 13 Graph of Neural Networks in GER

Insofar, no method has been able to eliminate overfitting. Bayesian method has proven that overcoming the above limitation is achievable. On top of that, the fact that a plethora of datasets are of small or average length generate another crucial point of concern that is solved by using the Bayesian method. A MAPE of 39% is not spectacular however the fact that this model follows correctly every up and down in the desired time horizon, is promising in on itself.

Below the predicted values are compared to the actual ones and each error contributes to the overall MAPE of this model (Table 6). The visual representation is also available for further consideration (Figure 14).

Table 6 Results of Bayesian in GER

	Actual Numbers	Predicted Numbers	Error
January	13385	10554.20	0.21
February	11140	9020.97	0.19
March	7935	8782.28	0.10
April	5645	8100.28	0.43
May	3775	8090.78	1.14
June	4790	8360.88	0.74

July	7590	9560.12	0.25
August	7275	9590.36	0.31
September	9300		0.05
		MAPE	**39%**

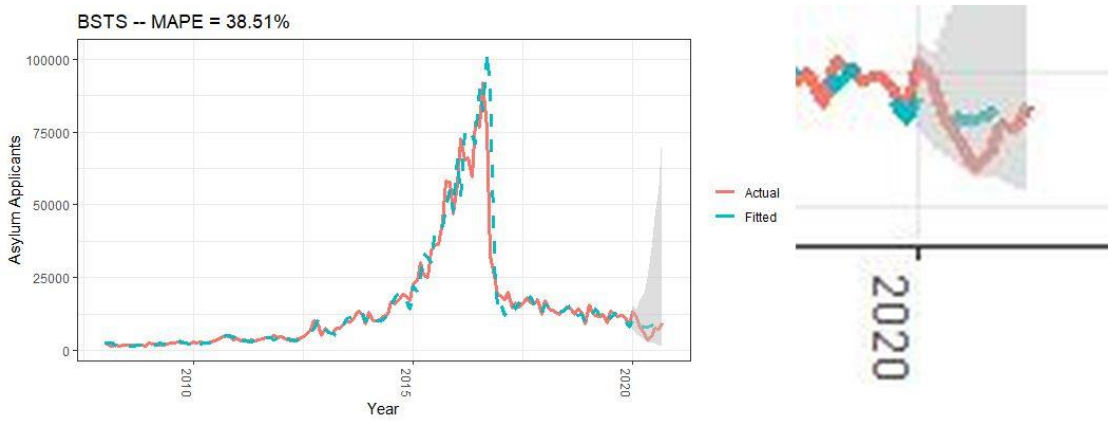


Figure 14 Graph of Bayesian method in GER

7.2 Italy

Italy is a European country that has been greatly affected by current affairs in the last decades. Firstly, its location makes it more likely for asylum-seekers to enter by sea. People of various countries of origin intend daily to find a shelter in Italy. It is not only a neighboring country to a lot of suffering communities, but also a developed humanitarian country. This has resulted in great increase of asylum claims. During 2020, the Coronavirus outbreak affected severely Italians to the point that strict measures regarding freedom of transport were implemented. All the above reasons explain the instability and erratic behavior of asylum-related migration phenomenon.

The same collection of models that was used in the case of Germany, was repurposed and used in the case of Italy. Unfortunately, the previous success that certain models had in predicting the time series behavior is nowhere to be found. The following graphs (Figures 15-22) are in the same order as above and the tables are missing because they couldn’t serve a purpose in such unstable sequences data.

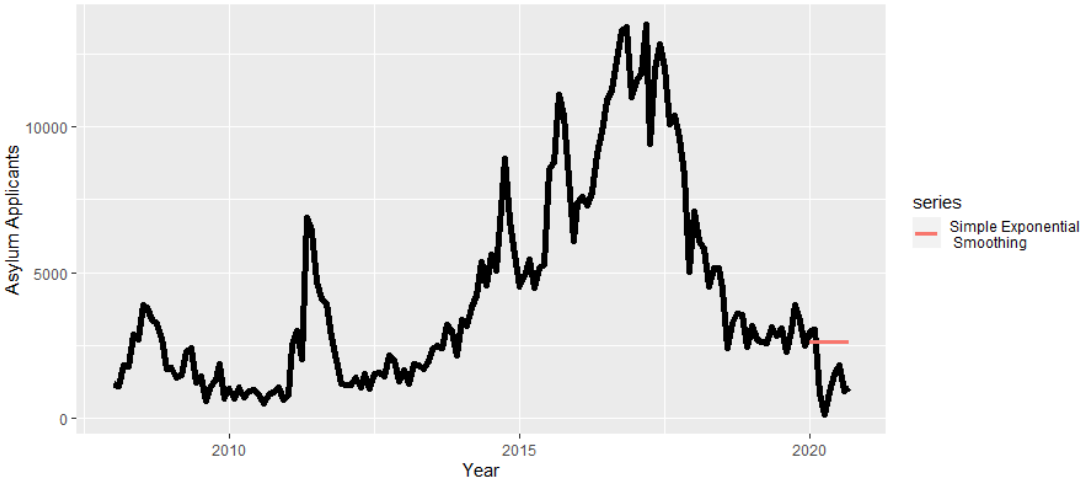


Figure 15 Graph of SES in ITL (MAPE: 327%)

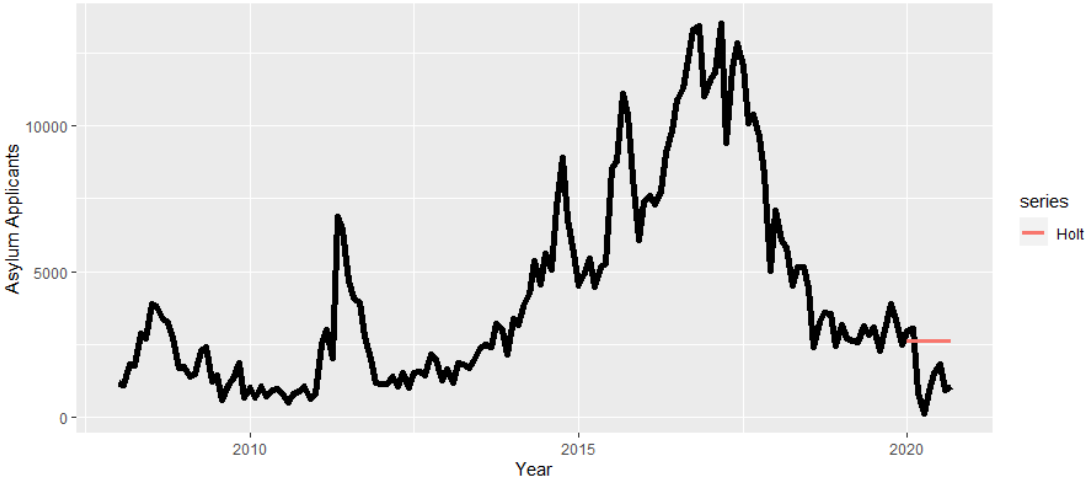


Figure 16 Graph of Holt in ITL (MAPE: 327%)

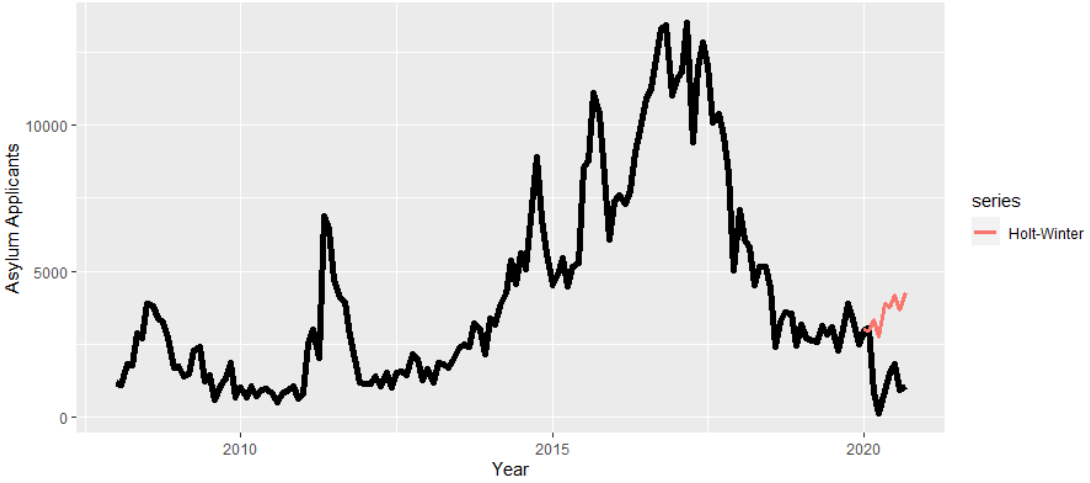


Figure 17 Graph of Holt-Winter in ITL (MAPE: 458%)

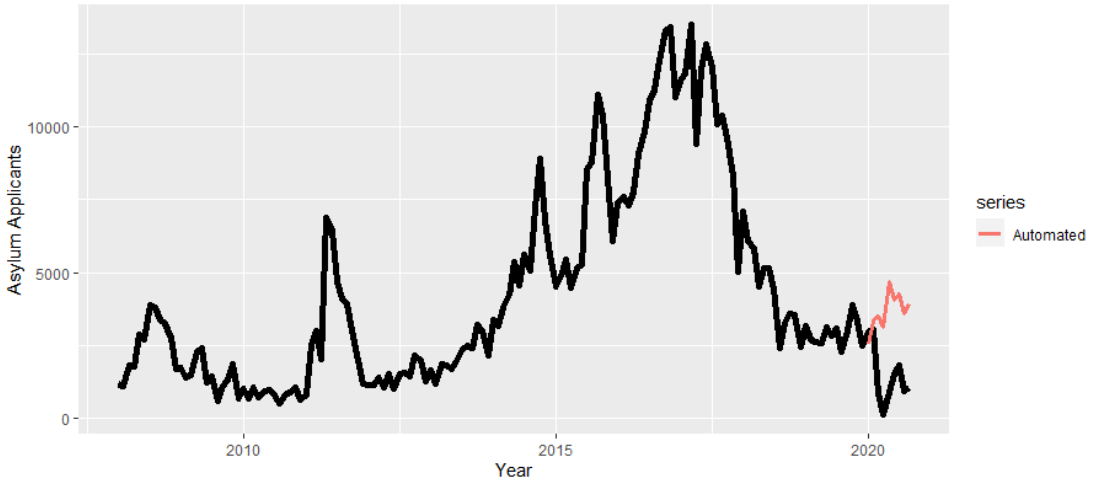


Figure 18 Graph of Automated in ITL(MAPE: 480%)

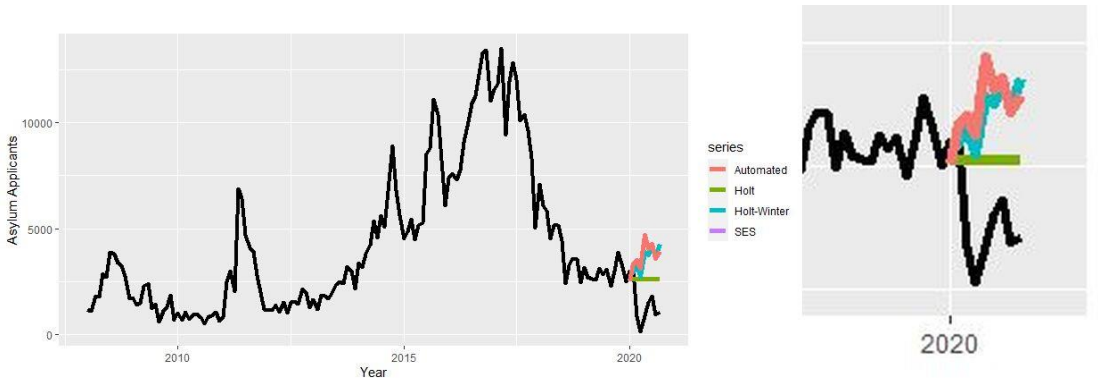


Figure 19 Graph of all Exponential Smoothing Models in ITL

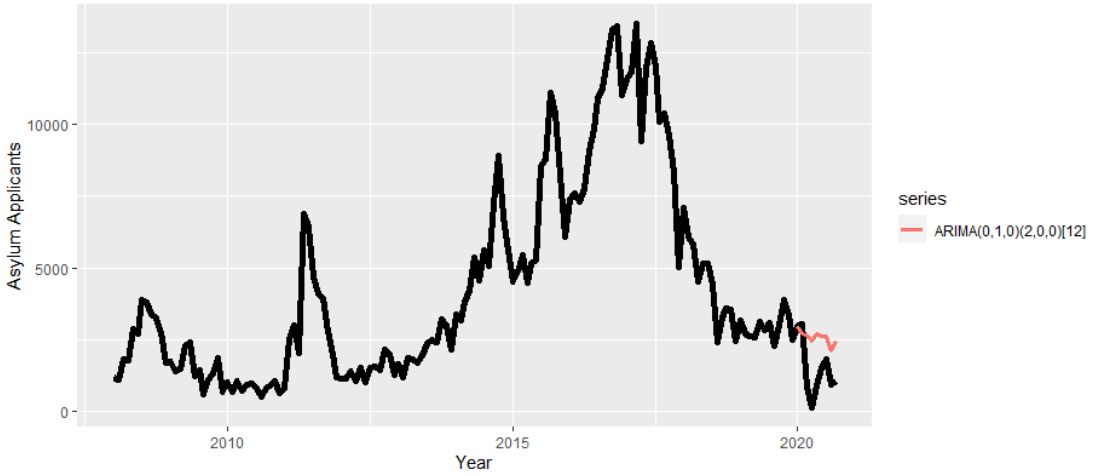


Figure 20 Graph of ARIMA(0,1,0)(2,0,0)[12] in ITL (MAPE: 305%)

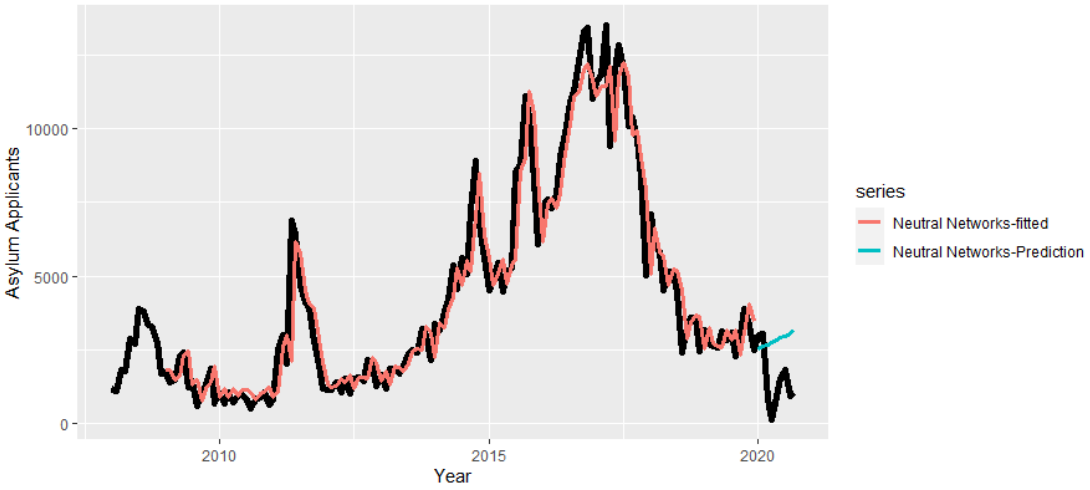


Figure 21 Graph of Neural Networks in ITL (MAPE: 359%)

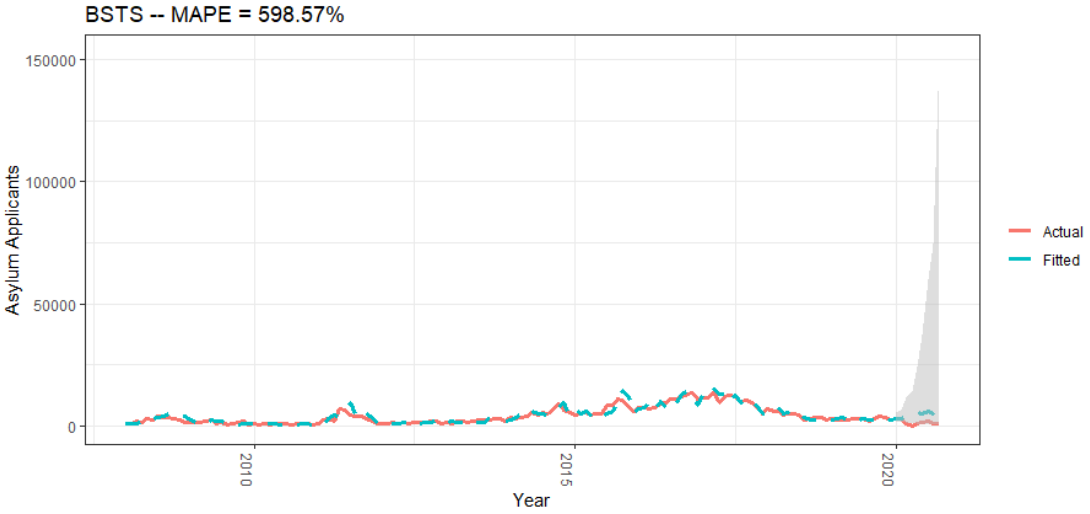


Figure 22 Graph of Bayesian in ITL (MAPE: 599%)

7.3 Norway

Scandinavian countries have been always an appealing destination for asylum-seekers. This primarily is the reason for selecting Norway as a third source of data. Norway is familiar and acknowledges the positives of a data-driven proactive approach towards population management. The graph that shows the annual time series of asylum claims in Norway is relatively stable apart from the sharp rise in the end of 2015 and the sharp drop in the beginning of 2016. The general trend is downward which continues all the way to the year 2020 when our test data are from.

The methodology remained unchanged in the case of Norway as well. Similarly to the case of Italy, none of the models were able to consistently predict realistic values. As a result, the method that produced the smallest MAPE is the Bayesian method with 68% error which is still above the set threshold. It is worth noting that no other model (Figures 23-29) produced results that could be considered even remotely trustworthy.

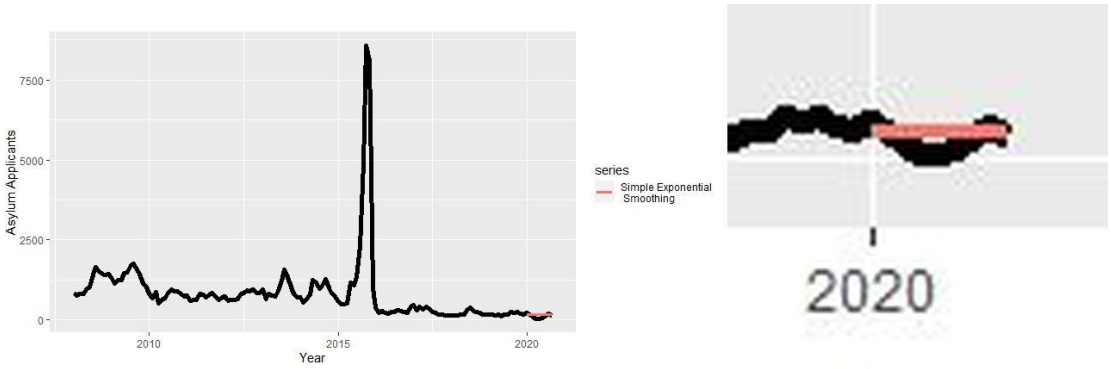


Figure 23 Graph of SES in NO (MAPE: 131%)

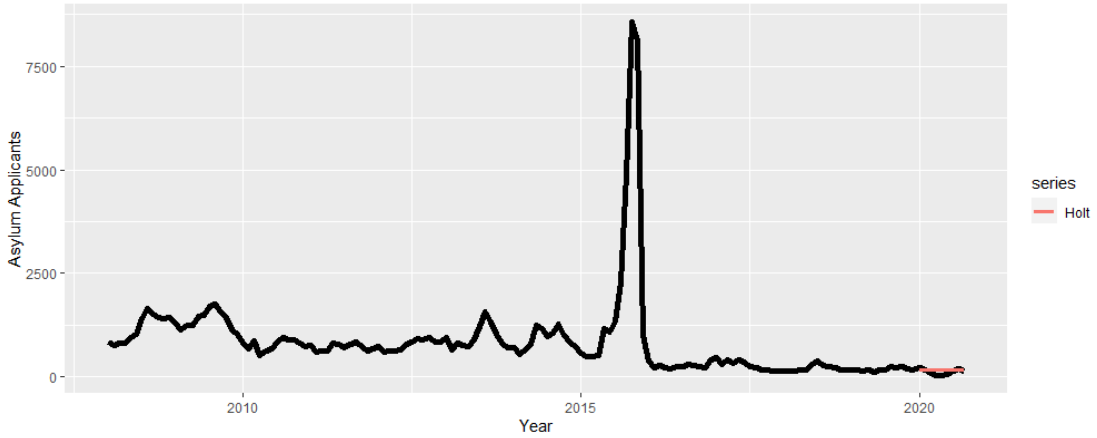


Figure 24 Graph of Holt in NO (MAPE: 131%)

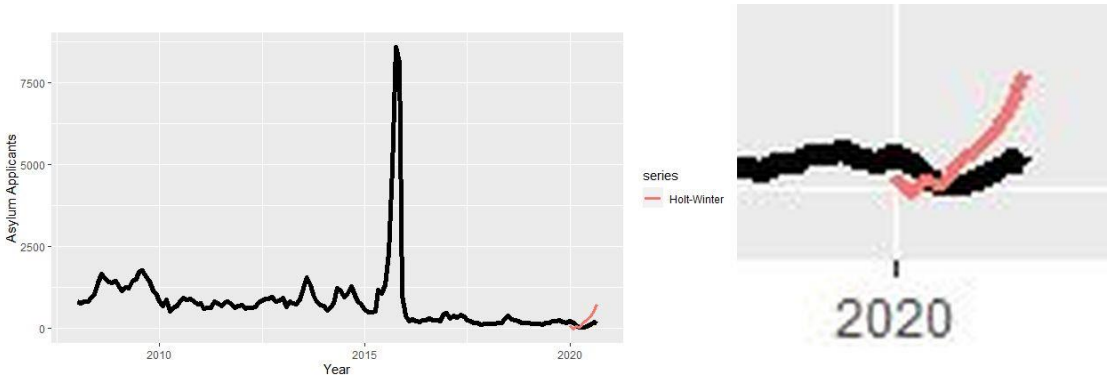


Figure 25 Graph of Holt-Winter in NO (MAPE: 225%)

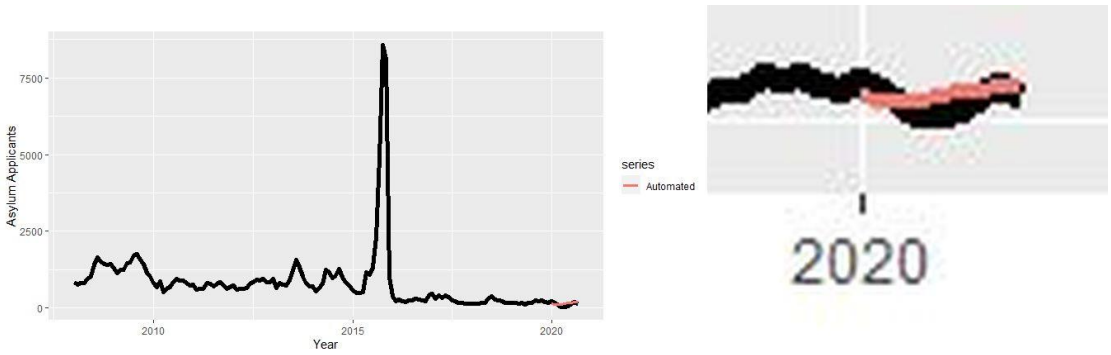


Figure 26 Graph of Automated in NO (MAPE: 110 %)

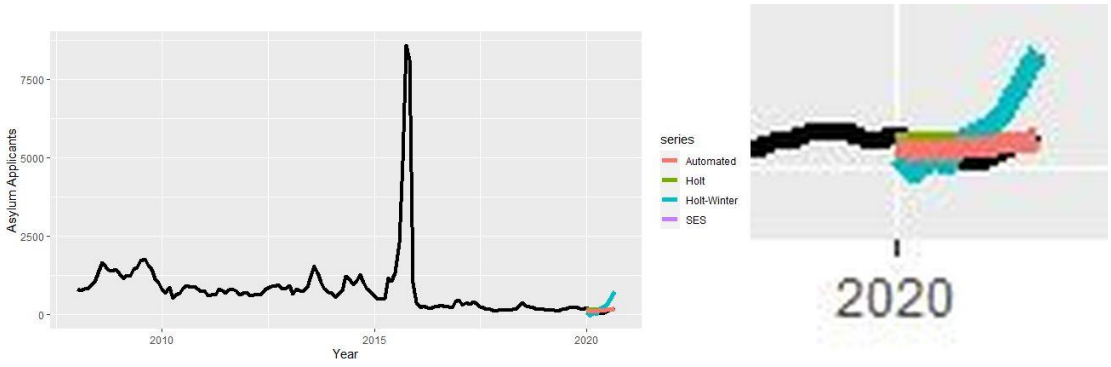


Figure 27 Graph of Exponential Smoothing Methods in NO

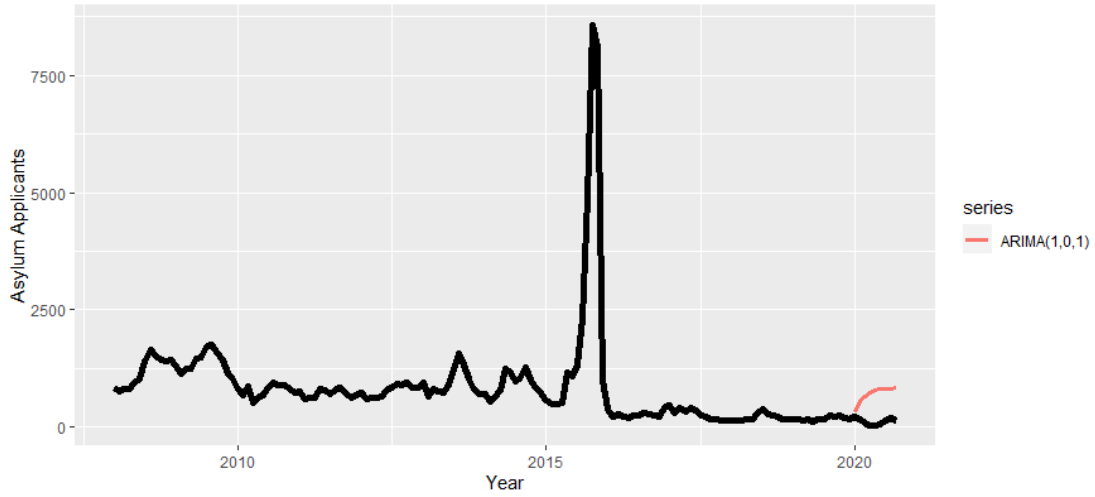


Figure 28 Graph of ARIMA(1,0,1) in NO (MAPE: 922%)

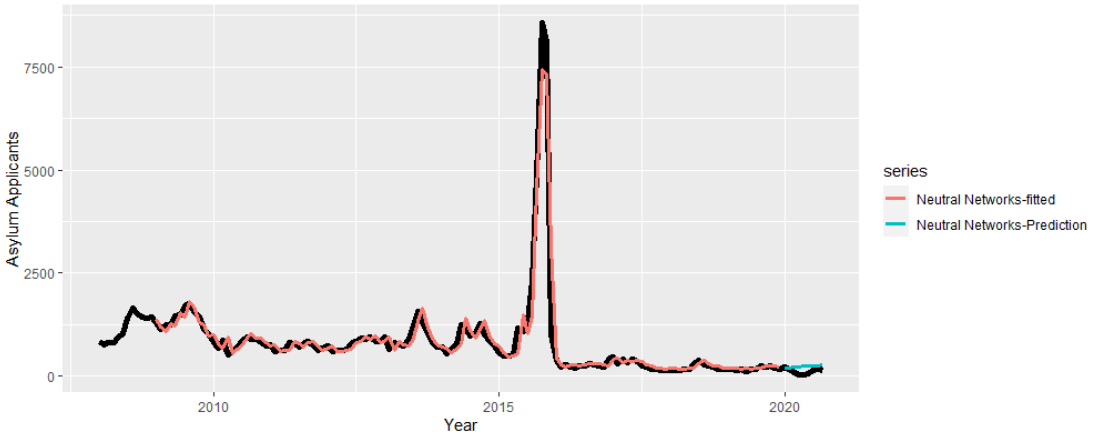
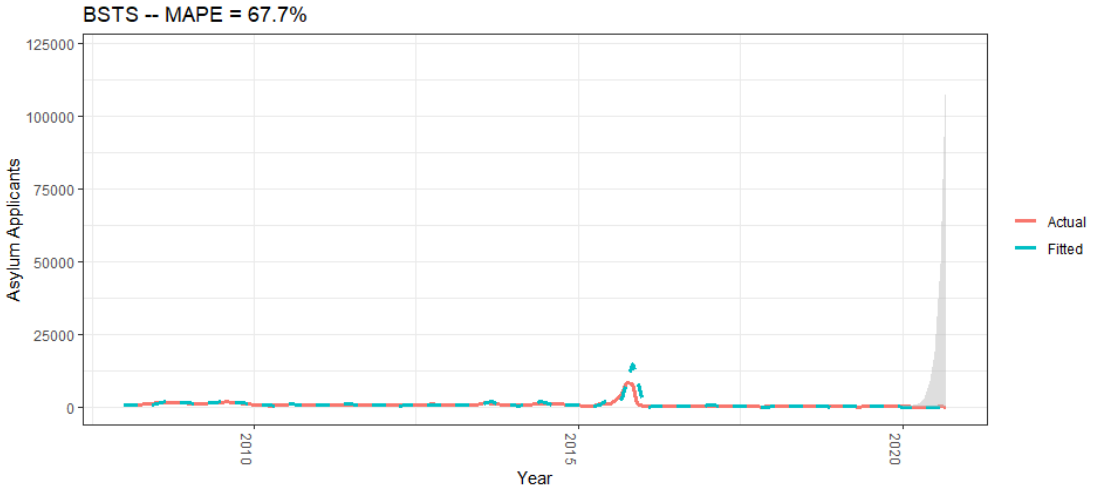


Figure 29 Graph of Neural Networks in NO (MAPE: 229%)



8 Discussion and Conclusion

Forced migration flows constitute the most volatile and sensitive type of migration. Over the years, refugees have been treated with suspicion from the media which has permeated into the various horizontal civic classes. Sources worldwide have been created to address those myths regarding the motives of such migration mobility. Some of those myths are that refugees are terrorists and extremists and that they pose a health hazard to locals. In 2015, the United Nations High Commissioner for Refugees Antonio Guterres proclaimed "*It is not the refugee outflows that cause terrorism, it is terrorism, tyranny and war that create refugees.*" Some other part of society believes that refugees

steal jobs from locals while undermining the local communities. The other side of the coin is that refugees are the most vulnerable group in the world and rely only on non-profit organizations to assist them with access to schooling, medical care and work in the country of destination. A recent phenomenon is refugees starting their own Refugee Learning Centers to provide guidance to new coming refugee children. It is important to highlight that a person is not considered a refugee under the Convention's definition just because they have fled a war zone. Applicants need to prove they have suffered or are at risk of suffering persecution.

Existing tools to monitor forced migration flows appear to be weak. Specifically, in this study almost all models failed to produce results under a specific error threshold. The reasons behind this fact have already been explained and they lead to the realization that only when many coexisting factors interact, they can influence so noticeably the results. Some European countries rely exclusively on statistical modelling. Although, models can be efficiently used as decision support systems, they are also widely used as early warning and alert systems. Some of the models are also used to fine tune those systems. This way the preparedness of policy makers is improved. Some of those systems are the EU Integrated Political Crisis Response (IPCR), the Displacement Tracking Matrix System (DTM) and the EASO's Early warning and Preparedness System. In the same vein, some online/offline surveys track people's intentions to migrate. The reality is that even with those supportive tools the forecasts produce big errors because the system's dynamic behavior cannot be tracked. An example of the current way a European country handles forecasting purposes is the Swedish model (Carammia & Dumont, 2018). Since 2018, the Swedish Migration Agency (SMA) uses forecasting models for short-term prediction. It is based on a Holt-Winters model (with seasonality) but it has to be adjusted several times a year to reflect every shock event.

A big part of the existing literature contains case studies where univariate time series were used as the main dataset. This significantly reduces the adaptability to varying situations. Forced migration can be the result of civil wars, poverty, human right violations and others. If all these factors are taken into account, a more holistic understanding of the intricate phenomenon will be acquired and lead to treat the causes of it and not its symptoms.

According to many sources the next big wave of forced migration will not be caused by an oppressive system or another pandemic but it will be from the climate change. Its

scale is to be orders of magnitude bigger than the Middle East crisis and this calls for the need of great predicting tools to be able to manage such a force of nature. This tremendous amount of people leaving due to environmental difficulties will be added to a system that already struggles in handling the current situation.

A valuable tool for figuring patterns of migration could be the information coming from search engines. Using words normally associated with migration, words like asylum regulations, refugee support funds etc. along with the country that has been found together could produce a signal that an individual is a potential asylum-seeker in the country of interest. For instance, if a great amount of searches from Venezuela include the terms asylum regulations in Germany, this might be a sign that a migration flow is expected to occur from Venezuela to Germany. After all, no country will be hurt for a slight overestimation. However, no one can promise that the opposite case is similarly painless.

9 Reference

Abel, G., Bijak, J., Findlay, A., McCollum, D., & Wiśniowski, A. (2013). Forecasting environmental migration to the United Kingdom: an exploration using Bayesian models. *Population and Environment*, *35*(2), 183-203.

Azose, J. J., & Raftery, A. E. (2013). Bayesian probabilistic projection of international migration rates. *arXiv preprint arXiv:1310.7148*.

Ahamer, G., Hardaker, G., Korotayev, A., & Zinkina, J. (2014). What does global migration network say about recent changes in the world system structure?. *Journal for Multicultural Education*.

Ahamer, G., Korotayev, A., & Zinkina, J. (2014). On the structure of the present-day convergence. *Campus-Wide Information Systems*.

Brown, R. G. (1959). *Statistical forecasting for inventory control*. McGraw/Hill

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797-817

Bakewell, O. (1999). Can we ever rely on refugee statistics?. *Radical Statistics*, *72*, 3-15.

Burstein, P. (2003). The impact of public opinion on public policy: A review and an agenda. *Political research quarterly*, *56*(1), 29-40.

Bijak, J. (2006, April). Forecasting international migration: Selected theories, models, and methods. Warsaw, Poland: Central European Forum for Migration Research.

Bijak, J., & Wiśniowski, A. (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *173*(4), 775-796.

Bassedas, P. M., Montijano, E. S., & i Lecha, E. S. (2015). *Europe and the Refugee Crisis: 10 Side-effects*. Barcelona Centre for International Affairs.

Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, *354*(6309), 217-222.

Brockwell P.J. and Davis R.A.,(2016), Introduction to time series and forecasting, Third edition, Springer

Chojnicki, Z. (1977). Podstawy metodologiczne prognozowania w geografii ekonomicznej [Methodological foundations of forecasting in economic geography]. Przegląd Geograficzny, 49(2): 247–261.

Castles, S. (2003). Towards a sociology of forced migration and social transformation. *sociology*, *37*(1), 13-34.

Chi, G., & Voss, P. (2005). Migration decision-making: a hierarchical regression approach. *Journal of Regional Analysis & Policy*, *35*(2), 11-22.

Castles, S., de Haas, H. and Miller, M.J. (2014), The Age of Migration: International Population Movements in the Modern World, 5th ed., The Guilford Press, New York, NY/London.

Carammia, M., & Dumont, J. C. (2018). Can we anticipate future migration flows. *Migration Policy Debates*, *OECD/EASO*, (16).

De Andrade, J. H. F. (2008). On the development of the concept of 'persecution'in international refugee law. *Brazilian Yearbook of International Law*, *3*(2), 114-136.

Docquier, F., Peri, G., & Ruyssen, I. (2014). The cross-country determinants of potential and actual migration. *International Migration Review*, *48*(1_suppl), 37-99

Disney, G., Wiśniowski, A., Forster, J. J., Smith, P. W., & Bijak, J. (2015). Evaluation of existing migration forecasting methods and models. *Report for the Migration Advisory Committee: Commissioned research. ESRC Centre for Population Change, University of Southampton*.

Douglas C. Montgomery, Cheryl L. Jennings, Murat Kulahci.(2015) Introduction to Time Series Analysis and Forecasting, , 2nd Edition, Wiley

EASO (2017) Quantitative assessment of asylum-related migration. A survey of methodology

Esipova, N., Pugliese, A., & Ray, J. (2018). More than 750 million worldwide would migrate if they could. *WORLD*, *13*, 14.

Holt, C. E. (1957). *Forecasting seasonals and trends by exponentially weighted averages* (O.N.R. Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, *60*(1), 126-142.

Hynes, T. (2003). The Issue of 'Trust'or 'Mistrust'in Research with Refugees: Choices. *Caveats and Considerations for Researchers, New Issues in Refugee Research Working Paper*, (98).

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

IOM (2020) THE FUTURE OF MIGRATION TO EUROPE: A systematic review of the literature on migration scenarios and forecasts

Kacapyr, E. (1996). *Economic forecasting: the state of the art*. ME Sharpe.

Kuschminder, K., de Bresser, J., & Siegel, M. (2015). Irregular migration routes to Europe and factors influencing migrants' destination choices. *Maastricht: Maastricht Graduate School of Governance*, 8-20.

Kick, E. L., & Davis, B. L. (2001). World-system structure and change: an analysis of global networks and economic growth across two time periods. *American Behavioral Scientist*, *44*(10), 1561-1578.

Keilman, N., Pham, D. Q., & Hetland, A. (2002). Why population forecasts should be probabilistic-illustrated by the case of Norway. *Demographic research*, *6*, 409-454.

Kupiszewski, M. (2002). How trustworthy are forecasts of international migration between Poland and the European Union? Journal of Ethnic and Migration Studies, 28(4): 627–645.

Keilman, N. (2007). UK national population projections in perspective: How successful compared to those in other European countries?. *POPULATION TRENDS-LONDON-*, *129*, 20.

Kalt, A., Hossain, M., Kiss, L., & Zimmerman, C. (2013). Asylum seekers, violence and health: a systematic review of research in high-income host countries. *American journal of public health*, *103*(3), e30-e42.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour.

Lee, E. S. (1966). A theory of migration. *Demography*, *3*(1), 47-57.

Ljones O. (2019). The importance of Official Refugee Statistics. Statistics Norway

Mahajan, V. (1976). The Delphi method: Techniques and applications. *JMR, Journal of Marketing Research (Pre-1986)*, *13*(000003), 317.

Mouzourakis, M. (2014). *'Wrong number?'The Use and Misuse of Asylum Data in the European Union*. CEPS Centre for European Policy Studies.

Öberg, S., and Wils, A. B. (1992). East-West migration in Europe. Can migration theories help estimate the numbers. Popnet, 22: 1–7.

Piore, M. J. (1979). *Birds of passage: Migrant labor and industrial societies*.

Portes, A. (1997). Immigration theory for a new century: Some problems and opportunities. *International migration review*, *31*(4), 799-825.

Prus, J., & Johnson, R. (1994). A critical review of student assessment options. *New Directions for Community Colleges*, *1994*(88), 69-83.

Ranis, G., & Fei, J. C. (1961). A theory of economic development. *The American economic review*, 533-565.

Robinson, V., & Segrott, J. (2002). *Understanding the decision-making of asylum seekers* (Vol. 12). London: Home Office.

Papastergiadis, N. (2004). The invasion complex in Australian political culture. *Thesis eleven*, *78*(1), 8-27.

Ruhrmann, H., & FitzGerald, D. (2017). The externalization of Europe's borders in the refugee crisis, 2015-2016.

Spinks, H. (2013). *Destination Anywhere?: Factors Affecting Asylum Seekers' Choice of Destination Country* (pp. 2012-13). Australia: Parliamentary Library.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of political Economy*, *70*(5, Part 2), 80-93.

Todaro, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American economic review*, *59*(1), 138-148.

UN (1949) Universal Declaration of Human Rights

UNHCR, (2010). Convention and Protocol Relating to the Status of Refugees

UNHCR (2015) Global Trends

UNHCR (2016) Global Trends

UNHCR (2017) Global Trends

UNHCR (2018) Global Trends

UNHCR (2019) Global Trends

UNHCR (2020) COVID-19 Emergency Measures in asylum and reception systems

Wallerstein, I. (1979). *The capitalist world-economy* (Vol. 2). Cambridge University Press.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, *6*, 324–342

Weible, C. M. (2008). Expert-based information and policy subsystems: a review and synthesis. *Policy Studies Journal*, *36*(4), 615-635.

Zudelova, T. Z. M. ECONOMETRIC MODELLING OF MIGRATION FLOWS.

Useful links:

UNHCR: https://www.unhcr.org
EASO: https://www.easo.europa.eu/
EUROSTAT: https://ec.europa.eu/eurostat
Asylum Information Database: http://www.asylumineurope.org/
Global Campaign for Equal Nationality Rights: https://www.equalnationalityrights.org/

10 Appendix

Trend, Seasonality and Randomness of every time series.

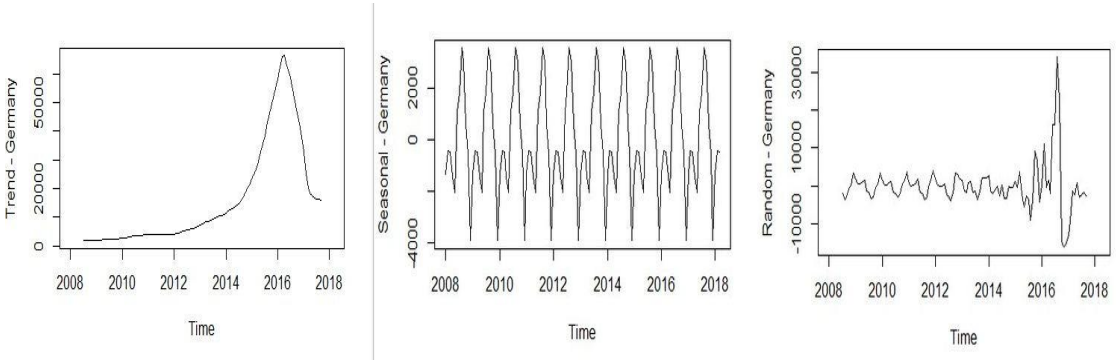


Figure 30 Germany

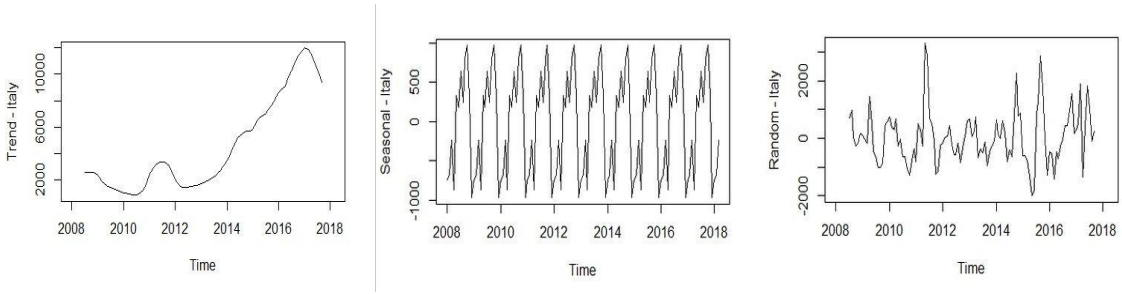


Figure 31 Italy

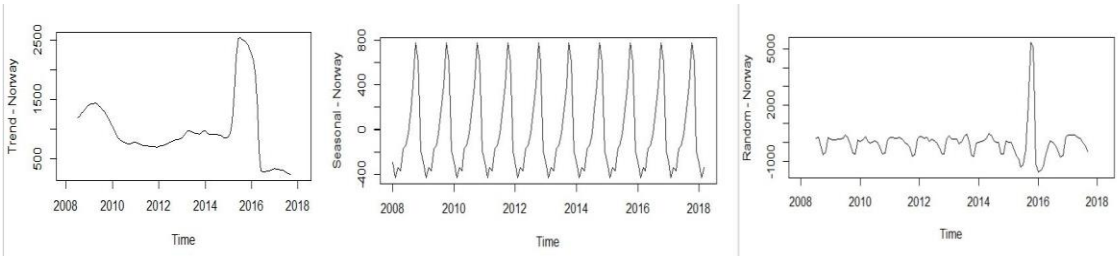


Figure 32 Norway

‘

Stationarity

Undoubtedly, different time series forecasting models are appropriate for different time series and in this section the conditions that should be met for each model will be explored.

The first criterium that needs to be checked and impacts completely the analytical and modeling tools used is stationarity. In a general term, stationarity is observed only in probabilistic (stochastic) processes and it means that the statistical properties of a process generating a time series do not change over time. In other words, the mean, variance and autocorrelation structure do not change over time. If any of those criteria is violated, the time series is not considered a stationary one. There are two forms of stationarity; strong stationarity and weak stationarity. There are tools and methods to ensure stationarity. The most popular time series forecasting methods for stationary time series are MA(p) or AR(q) and ARMA(p,q) models.

Moving Averages Models – MA(p)

Instead of using past values, a moving average model uses past forecast errors in a regression-like model. Thus, a moving average model of order q is always in the form

$$y_t = c + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q},$$

where εt is white noise sequence, hence moving average models are stationary since they are produced as a linear function of white noise units.

Auto-Regressive Models – AR(q)

An autoregressive model follows the same principles as a regression model with a difference. The independent variables or predictors are past values of the same variable. This is what the autoregression indicates. Thus, an autoregression model of order p is always in the form

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t,$$

Trend and Seasonality

The trend shows the general tendency of the data to increase or decrease during a long period of time. It can be stable as well. The most efficient fit of any systematic time pattern will produce the most accurate model. The tool that is used for the trend to be identified, is moving averages or regression analysis with the help of a trendline. A trendline can be linear or, as in most cases, nonlinear.

A trendline is an easily recognized line produced by connecting a series of data points. Its utility lies on its direction. Acknowledging its direction contributes to verifying if there is an upward or downward trend. They are so commonly used in the literature that they are topics of research obtaining their own terminology. Topics such as uptrends and downtrends have been popular topics many years around. The equation of a graph that contains a linear trendline is of the form

$$X_t = m_t + Y_t,$$

where mt is the trend component and Yt has zero mean.

Non-linear trend is when the pattern of the trendline is not a straight line and it cannot be expressed by the previous formula. There are numerous of formulas that are part of the non-linear trendline's collection. Polynomial, exponential and logarithmic trendlines are only some of the models of this category. A polynomial trendline is a curved line that is used when data fluctuates. The order of the polynomial can be determined by the number of fluctuations. An exponential trendline can be used only when positive values are observed and is most appropriate when data values rise or fall at increasingly higher rates. Lastly, a logarithmic trendline is a best-fit curved line when the rate of changes in the data increases or decreases quickly and then plateaus. Positive data values need to be present in this case-scenario as well.

Seasonality displayed in a time series can be put in other words as periodic fluctuations. Cycles that are observed to repeat over time such as monthly or yearly. Studying the seasonal component of a time series can improve the forecast predictability of the created model because a clearer relationship between input and output variables is presented and it can contribute to explanatory data analysis. There are time series that express complex seasonality, meaning that they may have a weekly pattern as well as an annual pattern. Data points with high frequency such as electricity consumption per minute are more likely to process complex seasonality. The forces that are unforeseen, un-

controllable, unpredictable and erratic constitute the last component; random or irregular movements.

EUROSTAT VS UNCHR

In order to examine whether the difference between the EUROSTAT records and UNCHR records of asylum-seekers is significantly important, inferential statistics and more specifically hypothesis testing is performed on them.

Ho: Their difference is significantly important

H1: Their difference is not significantly important

YEAR\AGENCY	EUROSTAT	UNCHR
2008	26845	22057
2009	32910	27631
2010	48480	41302
2011	53240	45717
2012	77485	54526
2013	126705	109558
2014	202645	173078
2015	476510	441899
2016	745160	722371
2017	222565	198301
2018	184180	133232
2019	165615	142493

Next, the samples' mean and standard deviation need to be calculated using the below formulas.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

	EUROSTAT	UNHCR
SAMPLE MEAN	196861,67	176013,75
ST. DEVIATION	212831,21	207294,42

Having as determinants that the two independent samples have unknown population variances but supposed to be the same, since they refer to the same country, the pooled variance formula is used. The pooled variance formula is always in the form

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$$

Next, the standard error of the pooled variance is needed using the below formula

$$s_p = \sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}$$

The sample size is 12 which is small and the population variances are unknown, hence the t-statistic will be used to compute the t-value.

$$t = \frac{(\overline{x_x} - \overline{x_y}) - 0}{st.pooled\ var} = 0{,}24$$

We cannot reject the Ho hypothesis so there is no sufficient data that there is any difference in the two different population means.

Active cases of persecution in 2020

Migration flows are highly sensitive to shock events, which creates a pervasive need for further research of the initial causes of the specific mobility which since late 2019 was also affected by the current Covid-19 outbreak effects. European Asylum Support Office (EASO), contains the most frequently updated list of asylum-related migration in Europe and it has its last entry in August of 2020. Scientists and academics endeavored to take into account the pandemic's influence on the migration crisis unsuccessfully, which is one of the core limitations of this phenomenon.

According to the United Nations High Commissioner on Refugees (UNHCR) in the advent of the Middle East crisis in 2015, 54% of all refugees worldwide came from just three countries. The Syrian Arab Republic, Afghanistan and Somalia. In 2016, the third position was replaced by South Sudan. This distribution was present until 2019, when this changed and the 3 countries that drove the majority of refugees were Syrian

Arab Republic, Venezuela and Afghanistan. The European countries that struggled the most were the ones located around the external borders of Europe and to the direction of the aforementioned countries of origin. In summary, Greece and Italy were afflicted mostly due to their unpreparedness.

The refugees generated by the above countries were spread worldwide. In 2015, when the middle east crisis started Europe received refuges mainly from Syrian Arab Republic, Afghanistan and Iraq and in actual numbers the primary destination countries were Germany, Hungary and Sweden. Nevertheless, population-wise Hungary, followed by Sweden and Austria were the top countries of destination with Germany to be the 6th country. What happened in Syrian Arab Republic, Afghanistan and Iraq and turned upside down the up to then "normality"?

Syrian Arab Republic

According to a summary of the crisis in Syria in BBC news, the crisis started in 2011 when some teenagers painted revolutionary slogans on a school wall. The brutal government response triggered nationwide protest demanding President Assad's resignation. Violence escalated and the country got into a civil war. The rise of the jihadist group Islamic State (IS) according to BBC added a further dimension in the case. Although President Assad agreed to remove chemical weapons, there are evidence that they were still used after the agreement, including some homemade chemical weapons. A UN commission of inquiry has evidence that all parties to the conflict have committed war crimes. Overall, in 2020, over 5.6 million people have fled Syria since 2011 to save themselves and their families.

Peace endeavors

Since 2011, there were more than 7 endeavors to reconvert Syria's prior state. Although all of them are important and worthy mention, some of them will be elaborated in this part just for the sake of completion.

In 2013 long initiatives and efforts for parties to work together around a full implementation of the Geneva Communique 2012 were settled. A report of BBC News presented a summary of the Geneva II conference, where UN, US, Russia and Syria's parties gathered to find a political solution. The fact that President Assad pursued this confer-

ence as an effort to eliminate terrorism declaring in advance that he was not planning to abdicate, combined with the reluctancy of the opposition party to negotiate unless President Assad stepped down, resulted in a conclusion that Geneva II would not be in a position to solve the crisis. After that Geneva III took place in 2016 and Geneva IV in 2017. In between, in 2015 the 'Four committees initiative' was proposed as a way to start a peace process in the Syrian Civil War. The last one took place in Kazakhstan in 2019 among Iran, Russia and Turkey when they expressed their unanimous stand by the Syrian President Assad and their motives to take care of internally displaced throughout Syria.

Afghanistan

According to BBC News, Afghan government and Taliban have been in a conflict for almost 20 years. Taliban militants' first took control in 1996. Their attack on 11 September 2001, however, fired the first US air strikes against Afghanistan. Since then, the Taliban were removed from power but US troops were still located in Afghanistan. During Obama's presidency American soldiers were calculated to be up to 100.000, the largest number until now. The number of NATO soldiers has been increasing and decreasing from time to time, but inevitably noncombatants have fled to neighboring countries, Asia and Europe.

Peace endeavors

The most profound agreement for peace was the one in the beginning of 2020, where Taliban militants signed a peace agreement with the United States. One of the requirements is that 5000 Taliban will have to set free, which is controlled by the Afghan government, hence it is not known if they will do so. The reality is that death toll after the US-Taliban agreement never drops down than 100 people monthly until September 2020. Next, Taliban and Afghan government need to agree on the terms and conditions of the beginning of a peace era.

Forced migration situations don't stop in the two above cases. Asylum-related applications by Venezuelan nationals have risen over 43000% since 2013. The most popular countries of destinations, however, are closer to Venezuela. Colombia, Peru, Chile,

Ecuador and Brazil being the top receiving countries. Iraq, Burundi, Nigeria and Sahel are some additional countries where forced displacement is a one-way street.